

XX The present invention relates to prophylaxis or treatment of infection
 CC by a microorganism in a biological environment comprising iron, heme
 CC or porphyrin. The treatment involves administering to the environment,
 CC an agent which antagonizes the interaction between a molecule derived
 CC from the microorganism and having an H42 domain, and an H42-binding
 CC motif on a porphyrin containing molecule present in the environment.
 CC Useful in the manufacture of a medicament for the prophylaxis and
 CC treatment of periodontal, pulmonary, vaginal, urethral or hoof disease
 CC resulting from *Porphyromonas gingivalis* infection or infection by a
 CC related microorganism.
 XX
 XX
 XX Sequence 21 AA:

Query Match	100.0%	Score 108;	DB 22;	Length 21;
Best Local Similarity	100.0%	Pred. No. 4.2e-12;		
Matches 21; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0

QY 1 ALNPDNYLSKDVATGATKKVY 21
| | | | | | | | | | | |
Db 1 ALNPDNYLSKDVATGATKKVY 21

RESULT 2
AAB49217
ID AAB49217 standard; protein; 134 AA

AC	AAB49217;
XX	
DT	13-MAR-2001 (first entry)

DE Peptide used in the invention.
XX
KW HA2 domain; porphyrin; periodontal; pulmonary; vaginal; urethral

OS Unidentified

PN WO2000072875-A1

PD 07-DEC-2000

26-MAY-2000; 2000WO-AU00599

PR 28-MAY-1999; 99AU-0000652

PA (UNSY) UNIV SYDNEY
XX

PI coll'yer CA, Hunter N, De Carlo AA
 NY

DR WPI; 2001-080424/09.
XX

PT Treating microbial infection in environment containing porphyrin, by
Administering a HA-2 antagonist -

PS Claim 9; Page 98-99; 102pp; English

The present invention relates to prophylaxis or treatment of infection caused by a microorganism in a biological environment comprising iron, heme or porphyrin. The treatment involves administering to the environment an agent which antagonizes the interaction between a molecule derived from the microorganism and having an HAZ domain, and an HAZ-binding motif on a porphyrin containing molecule present in the environment. Useful in the manufacture of a medicament for the prophylaxis and treatment of periodontal, pulmonary, vaginal, urethral or hoof diseases resulting from porphyromonas gingivalis infection or infection by a related microorganism.

Query Match	100.08;	Score 108;	DB 22;	Length 134;
Best Local Similarity	100.08;	Pred. No. 4.5e-11;		
Matches 21; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0

```

Oy      1 ALNPNDYLISKDVTGATKVKY 21
         ||| ||| ||| ||| ||| ||| |||
Db      60 ALNPNDYLISKDVTGATKVKY 80

```

RESULT 3
AAR96024

AC AAR96024

DT 04-SEP-1996 (first entry)

P. gingivalis haemagglutinin binds HAREP4 product

KW Haemagglutinin; haqA; periodontal disease; vaccine; antibody;
KW HAREDA

OS Porphyromonas gingivalis strain 381.

PN WQ9617936-A2

PD 13-JUN-1996.
xy

PF 11-DEC-1995; 95WO-US16108
XY

PR 09-DEC-1994; 94US-0353485.
XX

PA (UABR-) UAB RES FOUNDE
PA (LIVET.) INTV FLORIDA

XX	
PI	Han N., Iantz M., Levine G., patti JM., Proakis-Fox A

Pt
XX

Sumwasorn S;

DR WPL; 1990-28/161/29
DR N-PSDB; AAT30648.

PT porphyromonas gingivalis genes and proteins - used in the detection

XX F1 and vaccination against periodontal disease

STATE 4, PAGE 114-115, JOSEPH, EUGENIE.

of the *hagA* gene (AT30654) of *P. gingivalis* 318. It forms part

CC hayd repeat unit products (see also AAR96021-23) can be obtd. fr

CC animals against periodontal disease. Expression in *Salmonella*

CC to detect the presence of anti-P. gingivalis antibodies and to

XX

Query Match	Score	DB	Length
100.0%	108	17	439

Matches	21;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0
---------	-----	--------------	----	------------	----	--------	----	------	---

```
OY      1 ALNPDNYLISKDVTGATKVKY  21  
         |||||  
Db      251 ALNPDNYLISKDVTGATKVKY  271
```

RESULT 4
AAW69492
ID AAW69492 standard; Protein; 439 AA

```

DT      22-DEC-1998      (first entry)
XX
DE      Haemagglutinin protein ha9A, HA9ep4

```

```

XX Haemagglutinin protein; periodontal disease; vaccine; hga.
XX Porphyromonas gingivalis.
XX US5824791-A.
XX 20-OCT-1998.
XX 11-DEC-1995; 95US-0570311.
XX 11-DEC-1995; 95US-0570311.
XX 08-SEP-1988; 88US-0241640.
XX 25-JAN-1991; 91US-0647119.
XX 09-DEC-1994; 94US-0353485.
XX (UABR-) UAB RES FOUND.
XX (UFL) UNIV FLORIDA.
XX Han N, Lantz M, Lepine G, Patti JM, Progulsk-Fox A;
XX Tumwasorn S;
XX WPI: 1998-582627/49.
XX N-PSDB: AAV58879.
XX Isolated Porphyromonas gingivalis genes - encoding haemagglutinin
XX and/or protease poly:peptide(s))
XX Claim 1: Column 139-144; 101pp; English.
XX This sequence is encoded by a Porphyromonas gingivalis gene of the
XX invention. This sequence represents the hga haemagglutinin protein. The
XX polypeptides are used to produce antibodies to assist in the
XX periodontal disease. The antibodies are also used in purification and
XX identification procedures. The genes and polypeptides are used as
XX vaccines against periodontal disease.
XX Sequence 439 AA:
XX
XX Query Match 100.0%; Score 108; DB 19; Length 439;
XX Best Local Similarity 100.0%; Pred. No. 2e-10; Indels 0; Gaps 0;
XX Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 ALNPDNYLSKDYGTATKVKY 21
XX 251 ALNPDNYLSKDYGTATKVKI 271
XX
XX RESULT 5
XX AAR96021 standard; Protein: 450 AA.
XX AAR96021:
XX 04-SEP-1996 (first entry)
XX P. gingivalis haemagglutinin hga HAREPI product.
XX Haemagglutinin; hga; periodontal disease; vaccine; antibody;
XX HAREPI.
XX Porphyromonas gingivalis strain 381.
XX WO9617936-A2.
XX 13-JUN-1996.
XX 11-DEC-1995; 95MO-US16108.
XX 09-DEC-1994; 94US-0353485.
XX (UABR-) UAB RES FOUND.
XX (UFL) UNIV FLORIDA.

```

```

XX Han N, Lantz M, Lepine G, Patti JM, Progulsk-Fox A;
XX Tumwasorn S;
XX WPI: 1996-287181/29.
XX N-PSDB: AAT30645.
XX Porphyromonas gingivalis genes and proteins - used in the detection
XX and vaccination against periodontal disease
XX Claim 4: Page 103-104; 153pp; English.
XX HAREPI (AAR96021) is the product of the HAREPI repeat unit (AAT30645)
XX of the hga gene (AAT30654) of P. gingivalis 318. It forms part
XX of haemagglutinin hga (see also AAR96030). HAREPI and other
XX hga repeat unit products (see also AAR96022-24) can be obtd. from
XX transformed host cells and used as vaccines to protect humans or
XX animals against periodontal disease. Expression in Salmonella
XX cells allows produ. of live vaccine. HAREPI-4 can also be used
XX to detect the presence of anti-P. gingivalis antibodies and to
XX raise monoclonal antibodies for diagnostic appln.
XX Sequence 450 AA:
XX
XX Query Match 100.0%; Score 108; DB 17; Length 450;
XX Best Local Similarity 100.0%; Pred. No. 2.1e-10;
XX Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 ALNPDNYLSKDYGTATKVKY 21
XX 245 ALNPDNYLSKDYGTATKVKY 265
XX
XX RESULT 6
XX AAN69489 standard; Protein: 450 AA.
XX AAN69489:
XX 22-DEC-1998 (first entry)
XX Haemagglutinin protein hga, HAREPI.
XX Haemagglutinin protein; periodontal disease; vaccine; hga.
XX Porphyromonas gingivalis.
XX US5824791-A.
XX 20-OCT-1998.
XX 11-DEC-1995; 95US-0570311.
XX 11-DEC-1995; 95US-0570311.
XX 08-SEP-1988; 88US-0241640.
XX 25-JAN-1991; 91US-0647119.
XX 09-DEC-1994; 94US-0353485.
XX (UABR-) UAB RES FOUND.
XX (UFL) UNIV FLORIDA.
XX Han N, Lantz M, Lepine G, Patti JM, Progulsk-Fox A;
XX Tumwasorn S;
XX WPI: 1998-582627/49.
XX N-PSDB: AAV58876.
XX Isolated Porphyromonas gingivalis genes - encoding haemagglutinin
XX and/or protease poly:peptide(s))
XX Claim 1: Column 121-126; 101pp; English.
XX This sequence is encoded by a Porphyromonas gingivalis gene of the

```

CC Invention. This sequence represents the hga haemagglutinin protein. The
 CC polypeptides are used to produce antibodies to organisms associated with
 CC periodontal disease. The antibodies are also used in purification and
 CC identification procedures. The genes and polypeptides are used as
 CC vaccines against periodontal disease.

XX Sequence 450 AA:

Query Match 100.0%; Score 108; DB 19; Length 450;
 Best Local Similarity 100.0%; Pred. No. 2.1e-10;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALNPDNTLISKDVGTGATKVKY 21
 DB 245 ALNPDNTLISKDVGTGATKVKY 265

RESULT 7

AA96022 standard; Protein: 456 AA.

XX AAR96022:

XX 04-SEP-1996 (first entry)

XX P. gingivalis haemagglutinin hga Harep2 product.

XX Haemagglutinin: hga; periodontal disease; vaccine; antibody;

XX Harep2.

XX Porphyromonas gingivalis strain 381.

XX WO9617936-A2.

XX 13-JUN-1996.

XX 11-DEC-1995; 95WO-US16108.

XX 09-DEC-1994; 94US-0353485.

XX (UABR-) UAB RES FOUND.

XX (UFL) UNIV FLORIDA.

XX Han N, Lantz M, Lepline G, Patti JM, Progulsk-Fox A;

XX Tumwasorn S;

XX WPI: 1996-287181/29.

XX N-PSDB: AAT30646.

XX Porphyromonas gingivalis genes and proteins - used in the detection

XX and vaccination against periodontal disease

XX Claim 4: Page 107-108; 153pp: English.

XX Harep2 (AAR96022) is the product of the Harep2 repeat unit (AAT30646)

XX of the hga gene (AAT30654) of P. gingivalis 318. It forms part

XX of haemagglutinin hga (see also AAR96030). Harep2 and other hga

XX repeat unit products (see also AAR96021 and AAR96023-24) can be obtd.

XX from transformed host cells and used as vaccines to protect humans

XX or animals against periodontal disease. Expression in Salmonella

XX cells allows prodn. of live vaccine. Harep1-4 can also be used

XX to detect the presence of anti-P. gingivalis antibodies and to

XX raise monoclonal antibodies for diagnostic appln.

XX Sequence 456 AA:

Query Match 100.0%; Score 108; DB 17; Length 456;

Best Local Similarity 100.0%; Pred. No. 2.1e-10;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALNPDNTLISKDVGTGATKVKY 21

DB 251 ALNPDNTLISKDVGTGATKVKY 271

RESULT 8
 AAR96023 standard; Protein: 456 AA.

XX AAR96023:

XX 04-SEP-1996 (first entry)

XX P. gingivalis haemagglutinin hga Harep3 product.

XX Haemagglutinin: hga; periodontal disease; vaccine; antibody;

XX Harep3.

XX Porphyromonas gingivalis strain 381.

XX WO9617936-A2.

XX 13-JUN-1996.

XX 11-DEC-1995; 95WO-US16108.

XX 09-DEC-1994; 94US-0353485.

XX (UABR-) UAB RES FOUND.

XX (UFL) UNIV FLORIDA.

XX Han N, Lantz M, Lepline G, Patti JM, Progulsk-Fox A;

XX Tumwasorn S;

XX WPI: 1996-287181/29.

XX N-PSDB: AAT30647.

XX Porphyromonas gingivalis genes and proteins - used in the detection

XX and vaccination against periodontal disease

XX Claim 4: Page 110-112; 153pp: English.

XX Harep3 (AAR96023) is the product of the Harep3 repeat unit (AAT30647)

XX of the hga gene (AAT30654) of P. gingivalis 318. It forms part

XX of haemagglutinin hga (see also AAR96030). Harep3 and other hga

XX repeat unit products (see also AAR96021 and AAR96024) can be obtd.

XX from transformed host cells and used as vaccines to protect humans

XX or animals against periodontal disease. Expression in Salmonella

XX cells allows prodn. of live vaccine. Harep1-4 can also be used

XX to detect the presence of anti-P. gingivalis antibodies and to

XX raise monoclonal antibodies for diagnostic appln.

XX Sequence 456 AA:

Query Match 100.0%; Score 108; DB 17; Length 456;

Best Local Similarity 100.0%; Pred. No. 2.1e-10;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALNPDNTLISKDVGTGATKVKY 21

DB 251 ALNPDNTLISKDVGTGATKVKY 271

RESULT 9

AA96990 standard; Protein: 456 AA.

XX AA96990:

XX 22-DEC-1998 (first entry)

XX Haemagglutinin protein hga, Harep2.

XX Haemagglutinin protein: periodontal disease; vaccine; hga.

XX Porphyromonas gingivalis.


```

XX PN US5824791-A.
XX PD 20-OCT-1998.
XX PF 11-DEC-1995; 95US-0570311.
XX PR 11-DEC-1995; 95US-0570311.
XX PR 08-SEP-1988; 88US-0241640.
XX PR 25-JAN-1991; 91US-0647119.
XX PR 09-DEC-1994; 94US-0353485.
XX PA (UABR-) UAB RES FOUND.
XX PA (UYFL ) UNIV FLORIDA.
XX PI Han N, Lantz M, Lepine G, Patli JM, Proguliske-Fox A;
XX PI Tumwasorn S;
XX DR WPI: 1998-582627/49.
XX DR N-PSDB: AAV58877.
XX PT Isolated Porphyromonas gingivalis genes - encoding haemagglutinin
XX PT and/or protease poly(peptide(s))
XX PS Claim 1: Column 127-132; 101pp; English.
XX CC This sequence is encoded by a Porphyromonas gingivalis gene of the
XX CC invention. This sequence represents the hagd haemagglutinin protein. The
XX CC polypeptides are used to produce antibodies to organisms associated with
XX CC periodontal disease. The antibodies are also used in purification and
XX CC identification procedures. The genes and polypeptides are used as
XX CC vaccines against periodontal disease.
XX SQ Sequence 456 AA:

Query Match 100.0%; Score 108; DB 19; Length 456;
Best Local Similarity 100.0%; Pred. No. 2,1e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNPDNYLISKDVYTGATKVKY 21
DB 251 ALNPDNYLISKDVYTGATKVKY 271

RESULT 10
AAW69491
ID AAW69491 standard; Protein; 456 AA.
XX AC AAW69491;
XX DT 22-DEC-1998 (first entry)
XX DE Haemagglutinin protein hagd, HAREP3.
XX KW Haemagglutinin protein; periodontal disease; vaccine; hagd.
XX OS Porphyromonas gingivalis.
XX PN US5824791-A.
XX PD 20-OCT-1998.
XX PF 11-DEC-1995; 95US-0570311.
XX PR 11-DEC-1995; 95US-0570311.
XX PR 08-SEP-1988; 88US-0241640.
XX PR 25-JAN-1991; 91US-0647119.
XX PR 09-DEC-1994; 94US-0353485.
XX PA (UABR-) UAB RES FOUND.
XX PA (UYFL ) UNIV FLORIDA.
XX PI Han N, Lantz M, Lepine G, Patli JM, Proguliske-Fox A;

```

```

PI Tumwasorn S;
XX DR WPI: 1998-582627/49.
XX DR N-PSDB: AAV58878.
XX PT Isolated Porphyromonas gingivalis genes - encoding haemagglutinin
XX PT and/or protease poly(peptide(s))
XX PS Claim 1: Column 133-138; 101pp; English.
XX CC This sequence is encoded by a Porphyromonas gingivalis gene of the
XX CC invention. This sequence represents the hagd haemagglutinin protein. The
XX CC polypeptides are used to produce antibodies to organisms associated with
XX CC periodontal disease. The antibodies are also used in purification and
XX CC identification procedures. The genes and polypeptides are used as
XX CC vaccines against periodontal disease.
XX SQ Sequence 456 AA:

Query Match 100.0%; Score 108; DB 19; Length 456;
Best Local Similarity 100.0%; Pred. No. 2,1e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNPDNYLISKDVYTGATKVKY 21
DB 251 ALNPDNYLISKDVYTGATKVKY 271

RESULT 11
AAR96028
ID AAR96028 standard; Protein; 1087 AA.
XX AC AAR96028;
XX DT 04-SEP-1996 (first entry)
XX DE P. gingivalis haemagglutinin hagd.
XX KW Haemagglutinin; hagd; periodontal disease; vaccine; antibody.
XX OS Porphyromonas gingivalis strain FDC381.
XX PN W09617936-A2.
XX PD 13-JUN-1996.
XX PF 11-DEC-1995; 95WO-US16108.
XX PR 09-DEC-1994; 94US-0353485.
XX PA (UABR-) UAB RES FOUND.
XX PA (UYFL ) UNIV FLORIDA.
XX PI Han N, Lantz M, Lepine G, Patli JM, Proguliske-Fox A;
XX PI Tumwasorn S;
XX DR WPI: 1996-287181/29.
XX DR N-PSDB: AAT30652.
XX PT Porphyromonas gingivalis genes and proteins - used in the detection
XX PT and vaccination against periodontal disease
XX PS Claim 5; Page 65-68; 153pp; English.
XX CC P. gingivalis 381 haemagglutinin hagd (AAR96028) was identified as
XX CC the product of a gene (AAT30652) isolated from a P. gingivalis 318
XX CC genomic library. The haemagglutinin (see also AAR96032) can be obt.
XX CC from transformed host cells and used as a vaccine to protect humans
XX CC or animals against periodontal disease. Expression in Salmonella
XX CC cells allows prodn. of a live vaccine. The haemagglutinin can also
XX CC be used to detect the presence of anti-P. gingivalis antibodies and
XX CC to raise monoclonal antibodies for diagnostic apin.

```

SO Sequence 1087 AA:

Query Match 100.0%; Score 108; DB 17; Length 1087;
 Best Local Similarity 100.0%; Pred. No. 6.4e-10;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALNPDNYLISKDYTGATKVKY 21
 |||
 DB 579 ALNPDNYLISKDYTGATKVKY 599

RESULT 12
 AAW69486
 ID AAW69486 standard; Protein: 1087 AA.
 AC AAW69486;
 DT 22-DEC-1998 (first entry)
 DE Haemagglutinin protein haGD.
 KM Haemagglutinin protein: periodontal disease; vaccine; haGD.
 KN Porphyromonas gingivalis.
 OS US5824791-A.
 PN US5824791-A.
 PD 20-OCT-1998.
 PF 11-DEC-1995; 95US-0570311.
 PR 11-DEC-1995; 95US-0570311.
 PR 08-SEP-1988; 88US-0241640.
 PR 25-JAN-1991; 91US-0647119.
 PR 09-DEC-1994; 94US-0353485.
 PA (UABR-) UAB RES FOUND.
 PA (UYFL) UNIV FLORIDA.
 PI Han N, Lantz M, Lepine G, Patti JM, Progulskie-Fox A;
 PI Tummasorn S;
 DR WPI: 1998-582627/49.
 DR N-PSDB; AAV58873.
 PT Isolated Porphyromonas gingivalis genes - encoding haemagglutinin
 and/or protease poly(peptide(s))
 PS Claim 1; Column 57-64; 101pp; English.
 CC This sequence is encoded by a Porphyromonas gingivalis gene of the
 CC invention. This sequence represents the haGD haemagglutinin protein. The
 CC polypeptides are used to produce antibodies to organisms associated with
 CC periodontal disease. The antibodies are also used in purification and
 CC identification procedures. The genes and polypeptides are used as
 CC vaccines against periodontal disease.
 CC
 SQ Sequence 1087 AA:

Query Match 100.0%; Score 108; DB 19; Length 1087;
 Best Local Similarity 100.0%; Pred. No. 6.4e-10;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALNPDNYLISKDYTGATKVKY 21
 |||
 DB 579 ALNPDNYLISKDYTGATKVKY 599

RESULT 13
 AAR96032
 ID AAR96032 standard; Protein: 1358 AA.
 AC AAR96032;

XX 04-SEP-1996 (first entry)
 DT
 XX
 DE P. gingivalis haGD haemagglutinin.
 XX
 KM Haemagglutinin; haGD; periodontal disease; vaccine; antibody.
 XX
 OS Porphyromonas gingivalis strain FDC381.
 XX
 PN WO9617936-A2.
 PD 13-JUN-1996.
 PF 11-DEC-1995; 95WO-US16108.
 PR 09-DEC-1994; 94US-0353485.
 PA (UABR-) UAB RES FOUND.
 PA (UYFL) UNIV FLORIDA.
 PI Han N, Lantz M, Lepine G, Patti JM, Progulskie-Fox A;
 PI Tummasorn S;
 DR WPI: 1996-287181/29.
 DR N-PSDB; AAT30655.
 PT Porphyromonas gingivalis genes and proteins - used in the detection
 and vaccination against periodontal disease
 PS Claim 5; Page 125-129; 153pp; English.
 CC P. gingivalis 381 haemagglutinin haGD (AAR96032) was identified as
 CC the product of the second open reading frame of the haGD gene
 CC (AAT30655) derived from P. gingivalis 318 genomic DNA. A first
 CC open reading frame coded for haGD protease (see also AAR96031).
 CC The protease and haemagglutinin can be obtd. from transformed host
 CC cells and used in vaccines to protect humans or animals against
 CC periodontal disease. Expression in Salmonella cells allows prodn.
 CC of live vaccines. The haemagglutinin and protease can also be used
 CC to detect the presence of anti-P. gingivalis antibodies and to raise
 CC monoclonal antibodies for diagnostic apn.
 CC
 SQ Sequence 1358 AA:

Query Match 100.0%; Score 108; DB 17; Length 1358;
 Best Local Similarity 100.0%; Pred. No. 8.5e-10;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALNPDNYLISKDYTGATKVKY 21
 |||
 DB 850 ALNPDNYLISKDYTGATKVKY 870

RESULT 14
 AAW69494
 ID AAW69494 standard; Protein: 1358 AA.
 AC AAW69494;
 DT 22-DEC-1998 (first entry)
 DE Haemagglutinin protein haGD.
 KM Haemagglutinin protein: periodontal disease; vaccine; haGD.
 KN Porphyromonas gingivalis.
 OS US5824791-A.
 PN US5824791-A.
 PD 20-OCT-1998.
 PF 11-DEC-1995; 95US-0570311.

PR 11-DEC-1995: 95US-0570311.
 PR 08-SEP-1988: 88US-0241640.
 PR 25-JAN-1991: 91US-0647119.
 PR 09-DEC-1994: 94US-0353485.

XX (UABR-) UAB RES FOUND.
 PA (UYFL) UNIV FLORIDA.

XX Han N, Lantz M, Lepline G, Patti JM, Proguliske-Fox A;
 PI Tumwasorn S;

XX WPI: 1998-582627/49.
 DR N-PSDB: AAV58880.

XX Isolated Porphyromonas gingivalis genes - encoding haemagglutinin
 PT and/or protease poly(peptide(s))

XX Claim 1: Column 145-158; 101pp; English.

XX This sequence is encoded by a Porphyromonas gingivalis gene of the
 CC invention. This sequence represents the hagd haemagglutinin protein. The
 CC polypeptides are used to produce antibodies to organisms associated with
 CC periodontal disease. The antibodies are also used in purification and
 CC identification procedures. The genes and polypeptides are used as
 CC vaccines against periodontal disease.

XX Sequence 1358 AA;

Query Match 100.0%; Score 108; DB 19; Length 1358;
 Best Local Similarity 100.0%; Pred. No. 8.5e-10;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALNPDNYLISKDVTGATKVKY 21
 ||||||||||||||||||||
 DB 850 ALNPDNYLISKDVTGATKVKY 870

RESULT 15

AA96033
 ID AA96033 standard; Protein: 1687 AA.

XX AA96033;

XX 04-SEP-1996 (first entry)

XX P. gingivalis haemagglutinin hagd.

XX Haemagglutinin: hagd; periodontal disease; vaccine; antibody.

XX Porphyromonas gingivalis strain FDC381.

XX WO9617936-A2.

XX 13-JUN-1996.

XX 11-DEC-1995: 95WO-US16108.

XX 09-DEC-1994: 94US-0353485.

XX (UABR-) UAB RES FOUND.

XX (UYFL) UNIV FLORIDA.

XX Han N, Lantz M, Lepline G, Patti JM, Proguliske-Fox A;

XX Tumwasorn S;

XX WPI: 1996-287181/29.

XX N-PSDB: AAT30656.

XX Porphyromonas gingivalis genes and proteins - used in the detection
 PT and vaccination against periodontal disease
 PS Claim 5; Page 138-143; 153pp; English.
 XX

CC P. gingivalis 381 haemagglutinin hagd (AA96033) was identified as
 CC the product of a gene (AAT30656) identified in P. gingivalis 318
 CC genomic DNA. The haemagglutinin can be obtained from transformed
 CC host cells and used as a vaccine to protect humans or animals against
 CC periodontal disease. Expression in Salmonella cells allows production
 CC of a live vaccine. The haemagglutinin can also be used to detect the
 CC presence of anti-P. gingivalis antibodies and to raise monoclonal
 CC antibodies for diagnostic application.

XX Sequence 1687 AA;

Query Match 100.0%; Score 108; DB 17; Length 1687;
 Best Local Similarity 100.0%; Pred. No. 1.1e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALNPDNYLISKDVTGATKVKY 21
 ||||||||||||||||||||
 DB 1179 ALNPDNYLISKDVTGATKVKY 1199

Search completed: June 26, 2003, 00:51:42
 Job time: 7.41177 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 26, 2003, 00:50:09 ; Search time 2.20588 Seconds
(without alignments)
200.076 Million cell updates/sec

Title: US-09-980-370-1
Perfect score: 78
Sequence: 1 ALNPDNYLSKDVTC 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents-AA: *
1: /cgn2_6/prodata/1/1aa/5A.COMB.pep: *
2: /cgn2_6/prodata/1/1aa/5B.COMB.pep: *
3: /cgn2_6/prodata/1/1aa/6A.COMB.pep: *
4: /cgn2_6/prodata/1/1aa/6B.COMB.pep: *
5: /cgn2_6/prodata/1/1aa/PCTUS.COMB.pep: *
6: /cgn2_6/prodata/1/1aa/Backfilltest.pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	78	100.0	439	2	US-08-570-311-22
2	78	100.0	450	2	US-08-570-311-16
3	78	100.0	456	2	US-08-570-311-18
4	78	100.0	456	2	US-08-570-311-20
5	78	100.0	1087	2	US-08-570-311-8
6	78	100.0	1087	2	US-08-570-311-8
7	78	100.0	1358	2	US-08-570-311-27
8	78	100.0	1687	2	US-08-570-311-29
9	78	100.0	1704	3	US-08-336-308A-10
10	78	100.0	1704	3	US-08-822-324-6
11	78	100.0	1704	4	US-09-490-931-10
12	78	100.0	1732	2	US-08-570-311-10
13	78	100.0	1732	2	US-08-353-485-10
14	78	100.0	2628	2	US-08-570-311-14
15	46	59.0	497	2	US-08-570-311-2
16	46	59.0	497	2	US-08-353-485-2
17	42	53.8	2548	4	US-09-172-422-1
18	40	51.3	248	1	US-08-152-019A-14
19	40	51.3	249	1	US-08-152-019A-35
20	40	51.3	249	1	US-08-144-121-6
21	40	51.3	249	2	US-08-460-309-6
22	40	51.3	249	2	US-08-125-077-6
23	40	51.3	249	2	US-08-735-893-6
24	40	51.3	1196	1	US-08-144-121-4
25	40	51.3	1196	1	US-08-735-893-4
26	39	50.0	331	2	US-08-997-080-182
27	39	50.0	331	2	US-08-997-362-182

28	39	50.0	331	4	US-09-095-855-182	Sequence 182, App
29	39	50.0	331	4	US-09-324-542-182	Sequence 182, App
30	39	50.0	331	4	US-09-205-426-182	Sequence 182, App
31	38	48.7	638	2	US-08-846-762-95	Sequence 95, App
32	37	47.4	275	1	US-07-779-890-8	Sequence 8, App
33	37	47.4	275	1	US-07-779-890-8	Sequence 8, App
34	37	47.4	275	5	PCT-US93-05640-8	Sequence 8, App
35	37	47.4	416	4	US-09-228-246-4	Sequence 4, App
36	37	47.4	509	1	US-07-779-890-6	Sequence 6, App
37	37	47.4	509	1	US-07-779-890-6	Sequence 6, App
38	37	47.4	509	2	US-09-008-962-3	Sequence 3, App
39	37	47.4	509	2	US-08-673-307-3	Sequence 3, App
40	37	47.4	509	3	US-09-213-205-3	Sequence 3, App
41	37	47.4	509	3	US-08-733-360A-10	Sequence 10, App
42	37	47.4	509	4	US-08-916-935-11	Sequence 11, App
43	37	47.4	509	5	PCT-US93-05640-6	Sequence 6, App
44	37	47.4	524	1	US-08-447-500-24	Sequence 24, App
45	37	47.4	524	1	US-08-454-097-24	Sequence 24, App

ALIGNMENTS

RESULT 1
US-08-570-311-22
Sequence 22, Application US/08570311
Patent No. 5824791
GENERAL INFORMATION:
APPLICANT: Proguiske-Fox, Ann
APPLICANT: Tumwasorn, Somying
APPLICANT: Lepine, Guylaine
APPLICANT: Han, Naiming
APPLICANT: Lantz, Marilyn
APPLICANT: Patil, Joseph
TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
and Probes for the Detection of Periodontal Disease
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ted W. Whitlock
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/570,311
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/353,485
FILING DATE: 09-DEC-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/647,119
FILING DATE: 25-JAN-1991
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/241,640
FILING DATE: 08-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: 0F15.C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:

LENGTH: 439 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-570-311-22

Query Match 100.0%; Score 78; DB 2; Length 439;
Best Local Similarity 100.0%; Pred. No. 6.4e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALNPDNYLISKDVTG 15
|||||
DB 251 ALNPDNYLISKDVTG 265

RESULT 2
US-08-570-311-16
Sequence 16, Application US/08570311
Patent No. 5824791

GENERAL INFORMATION:
APPLICANT: Proguiske-Fox, Ann
APPLICANT: Tumwasorn, Somying
APPLICANT: Lepine, Guylaine
APPLICANT: Han, Naiping
APPLICANT: Lantz, Marilyn
TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ted W. Whitlock
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/570,311
FILING DATE: 25-JAN-1991

CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/353,485
FILING DATE: 09-DEC-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/647,119
FILING DATE: 25-JAN-1991
CLASSIFICATION: 424

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/241,640
FILING DATE: 08-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UF15.C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 450 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-570-311-16

Query Match 100.0%; Score 78; DB 2; Length 450;
Best Local Similarity 100.0%; Pred. No. 6.6e-06;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 ALNPDNYLISKDVTG 15
|||||
DB 245 ALNPDNYLISKDVTG 259

RESULT 3
US-08-570-311-18
Sequence 18, Application US/08570311
Patent No. 5824791

GENERAL INFORMATION:
APPLICANT: Proguiske-Fox, Ann
APPLICANT: Tumwasorn, Somying
APPLICANT: Lepine, Guylaine
APPLICANT: Han, Naiping
APPLICANT: Lantz, Marilyn
TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ted W. Whitlock
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/570,311
FILING DATE: 25-JAN-1991

CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/353,485
FILING DATE: 09-DEC-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/647,119
FILING DATE: 25-JAN-1991
CLASSIFICATION: 424

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/241,640
FILING DATE: 08-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UF15.C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 456 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-570-311-18

Query Match 100.0%; Score 78; DB 2; Length 456;
Best Local Similarity 100.0%; Pred. No. 6.7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 ALNPDNYLISKDVTG 15
|||||
DB 251 ALNPDNYLISKDVTG 265

RESULT 4

US-08-570-311-20

Sequence 20, Application US/08570311

Patent No. 5824791

GENERAL INFORMATION:

APPLICANT: Proguiske-Fox, Ann

APPLICANT: Tumwasorn, Somying

APPLICANT: Lepine, Guylaine

APPLICANT: Han, Naïming

APPLICANT: Lantz, Marilyn

APPLICANT: Patti, Joseph

TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes

TITLE OF INVENTION: and probes for the Detection of Periodontal Disease

NUMBER OF SEQUENCES: 29

CORRESPONDENCE ADDRESS:

ADDRESSEE: Ted W. Whitlock

STREET: 2421 N.W. 41st Street, Suite A-1

CITY: Gainesville

STATE: FL

COUNTRY: USA

ZIP: 32606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/570,311

FILING DATE:

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/353,485

FILING DATE: 09-DEC-1994

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/647,119

FILING DATE: 25-JAN-1991

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/241,640

FILING DATE: 08-SEP-1988

ATTORNEY/AGENT INFORMATION:

NAME: Whitlock, Ted W.

REGISTRATION NUMBER: 36,965

REFERENCE/DOCKET NUMBER: UF15.C3

TELEPHONE: (904) 375-8100

TELEFAX: (904) 372-5800

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:

LENGTH: 456 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-570-311-20

Query Match 100.0%; Score 78; DB 2; Length 456;
Best Local Similarity 100.0%; Pred. No. 6.7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALNPDNYLSKDYTG 15

DB 251 ALNPDNYLSKDYTG 265

RESULT 5

US-08-570-311-8

Sequence 8, Application US/08570311

Patent No. 5824791

GENERAL INFORMATION:

APPLICANT: Proguiske-Fox, Ann

APPLICANT: Tumwasorn, Somying

APPLICANT: Lepine, Guylaine

APPLICANT: Han, Naïming

APPLICANT: Lantz, Marilyn
APPLICANT: Patti, Joseph
TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
TITLE OF INVENTION: and probes for the Detection of Periodontal Disease
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:

ADDRESSEE: Ted W. Whitlock

STREET: 2421 N.W. 41st Street, Suite A-1

CITY: Gainesville

STATE: FL

COUNTRY: USA

ZIP: 32606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/570,311

FILING DATE:

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/353,485

FILING DATE: 09-DEC-1994

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/647,119

FILING DATE: 25-JAN-1991

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/241,640

FILING DATE: 08-SEP-1988

ATTORNEY/AGENT INFORMATION:

NAME: Whitlock, Ted W.

REGISTRATION NUMBER: 36,965

REFERENCE/DOCKET NUMBER: UF15.C3

TELEPHONE: (904) 375-8100

TELEFAX: (904) 372-5800

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 1087 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-570-311-8

Query Match 100.0%; Score 78; DB 2; Length 1087;
Best Local Similarity 100.0%; Pred. No. 1.9e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALNPDNYLSKDYTG 15

DB 579 ALNPDNYLSKDYTG 593

RESULT 6

US-08-353-485-8

Sequence 8, Application US/08353485

Patent No. 5830710

GENERAL INFORMATION:

APPLICANT: Proguiske-Fox, Ann

APPLICANT: Tumwasorn, Somying

APPLICANT: Lepine, Guylaine

APPLICANT: Han, Naïming

APPLICANT: Lantz, Marilyn

APPLICANT: Patti, Joseph

TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes

TITLE OF INVENTION: and probes for the Detection of Periodontal Disease

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Ted W. Whitlock

STREET: 2421 N.W. 41st Street, Suite A-1

CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/353,485
FILING DATE: 09-DEC-1994
CLASSIFICATION: 424
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/647,119
FILING DATE: 25-JAN-1991
CLASSIFICATION: 424
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/241,640
FILING DATE: 08-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UFI5.C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1087 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-353-485-8

Query Match 100.0%; Score 78; DB 2; Length 1087;
Best Local Similarity 100.0%; Pred. No. 1.9e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNPDNYLISKDVVG 15
DB 579 ALNPDNYLISKDVVG 593

RESULT 7
US-08-570-311-27
Sequence 27, Application US/08570311
Patent No. 5824791
GENERAL INFORMATION:
APPLICANT: Proguiske-Fox, Ann
APPLICANT: Tumwasorn, Somying
APPLICANT: Lepine, Guylaine
APPLICANT: Han, Namling
APPLICANT: Lantz, Marilyn
APPLICANT: Patti, Joseph
TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ted W. Whitlock
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/570,311
FILING DATE:

CLASSIFICATION: 424
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/353,485
FILING DATE: 09-DEC-1994
CLASSIFICATION: 424
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/647,119
FILING DATE: 25-JAN-1991
CLASSIFICATION: 424
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/241,640
FILING DATE: 08-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UFI5.C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 1358 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-570-311-27

Query Match 100.0%; Score 78; DB 2; Length 1358;
Best Local Similarity 100.0%; Pred. No. 2.5e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNPDNYLISKDVVG 15
DB 850 ALNPDNYLISKDVVG 864

RESULT 8
US-08-570-311-29
Sequence 29, Application US/08570311
Patent No. 5824791
GENERAL INFORMATION:
APPLICANT: Proguiske-Fox, Ann
APPLICANT: Tumwasorn, Somying
APPLICANT: Lepine, Guylaine
APPLICANT: Han, Namling
APPLICANT: Lantz, Marilyn
APPLICANT: Patti, Joseph
TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ted W. Whitlock
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/570,311
FILING DATE:
CLASSIFICATION: 424
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/353,485
FILING DATE: 09-DEC-1994
CLASSIFICATION: 424
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/647,119
FILING DATE: 25-JAN-1991

CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/241,640
FILING DATE: 08-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: DF15-C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 1687 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-570-311-29

Query Match 100.0%; Score 78; DB 2; Length 1687;
Best Local Similarity 100.0%; Pred. No. 3.2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALPNPNYLISKDVYG 15
|||||

Db 1179 ALPNPNYLISKDVYG 1193

RESULT 9
US-08-336-308A-10
Sequence 10, Application US/08336308A
Patent No. 6017532
GENERAL INFORMATION:
APPLICANT: Travls, James
APPLICANT: Potempa, Jan S.
APPLICANT: Bart, Philip J.
APPLICANT: Pavloff, Nadine
TITLE OF INVENTION: Porphyromonas gingivalis
TITLE OF INVENTION: Arginine-specific Proteinase Coding Sequences
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: US
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/336,308A
FILING DATE: 08-NOV-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/119,361
FILING DATE: 10-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/265,441
FILING DATE: 24-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Ferber, Donna M.
REGISTRATION NUMBER: 33,878
REFERENCE/DOCKET NUMBER: 21-93C
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 499-8080
TELEFAX: (303) 499-8089
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1704 amino acids
TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-336-308A-10

Query Match 100.0%; Score 78; DB 3; Length 1704;
Best Local Similarity 100.0%; Pred. No. 3.3e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALPNPNYLISKDVYG 15
|||||

Db 1196 ALPNPNYLISKDVYG 1210

RESULT 10
US-08-822-324-6
Sequence 6, Application US/08822324
Patent No. 6129917
GENERAL INFORMATION:
APPLICANT: Potempa, Jan S.
APPLICANT: Travls, James
APPLICANT: Genco, Caroline A.
TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS COMPRISING
TITLE OF INVENTION: PORPHYROMONAS GINGIVALIS PROTEINS AND/OR PEPTIDES AND
TITLE OF INVENTION: METHODS
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: CO
COUNTRY: US
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/822,324
FILING DATE: 21-MAR-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/013,945
FILING DATE: 22-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ferber, Donna M.
REGISTRATION NUMBER: 33,878
REFERENCE/DOCKET NUMBER: 103-95 WO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 488-8080
TELEFAX: (303) 499-8089
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1704 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-822-324-6

Query Match 100.0%; Score 78; DB 3; Length 1704;
Best Local Similarity 100.0%; Pred. No. 3.3e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALPNPNYLISKDVYG 15
|||||

Db 1196 ALPNPNYLISKDVYG 1210

RESULT 11
US-09-490-931-10
Sequence 10, Application US/09490931
Patent No. 6274718
GENERAL INFORMATION:

APPLICANT: Travis, James
APPLICANT: Potempa, Jan S.
APPLICANT: Barr, Philip J.
APPLICANT: Pavloff, Nadine
TITLE OF INVENTION: Porphyromonas gingivalis
TITLE OF INVENTION: Arginine-specific Proteinase Coding Sequences
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: US
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/490,931
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/336,308
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/265,441
FILING DATE: 24-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Felder, Donna M.
REGISTRATION NUMBER: 33,878
REFERENCE/DOCKET NUMBER: 21-93C
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 499-8080
TELEFAX: (303) 499-8089
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1704 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-490-931-10
Query Match 100.0%; Score 78; DB 4; Length 1704;
Best Local Similarity 100.0%; Pred. No. 3.3e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 ALNPDNYLSKDVGTG 15
DB 1196 ALNPDNYLSKDVGTG 1210
RESULT 12
US-08-570-311-10
Sequence 10, Application US/08570311
Patent No. 5824791
GENERAL INFORMATION:
APPLICANT: Proguiske-Fox, Ann
APPLICANT: Tumwasorn, Somying
APPLICANT: Lepine, Guyalaine
APPLICANT: Han, NaIning
APPLICANT: Lantz, Marilyn
APPLICANT: Patti, Joseph
TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ted W. Whitlock
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA

ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/570,311
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/353,485
FILING DATE: 09-DEC-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/647,119
FILING DATE: 25-JAN-1991
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/241,640
FILING DATE: 08-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UF15.C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1732 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-570-311-10
Query Match 100.0%; Score 78; DB 2; Length 1732;
Best Local Similarity 100.0%; Pred. No. 3.3e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 ALNPDNYLSKDVGTG 15
DB 1216 ALNPDNYLSKDVGTG 1230
RESULT 13
US-08-353-485-10
Sequence 10, Application US/08353485
Patent No. 5830710
GENERAL INFORMATION:
APPLICANT: Proguiske-Fox, Ann
APPLICANT: Tumwasorn, Somying
APPLICANT: Lepine, Guyalaine
APPLICANT: Han, NaIning
APPLICANT: Lantz, Marilyn
APPLICANT: Patti, Joseph
TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ted W. Whitlock
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/353,485

;; FILING DATE: 09-DEC-1994
;; CLASSIFICATION: 424
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/647,119
;; FILING DATE: 25-JAN-1991
;; CLASSIFICATION: 424
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/241,640
;; FILING DATE: 08-SEP-1988
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Whitlock, Ted W.
;; REGISTRATION NUMBER: 36,965
;; REFERENCE/DOCKET NUMBER: UF15.C2
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (904) 375-8100
;; TELEFAX: (904) 372-5800
;; INFORMATION FOR SEQ ID NO: 10:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1732 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-353-485-10

Query Match 100.0%; Score 78; DB 2; Length 1732;
Best Local Similarity 100.0%; Pred. No. 3.3e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ALNPDNYLSKDVGT 15
Db 1216 ALNPDNYLSKDVGT 1230

RESULT 14
US-08-570-311-14
; Sequence 14, Application US/08570311
; Patent No. 5824791
; GENERAL INFORMATION:
; APPLICANT: Proguiske-Fox, Ann
; APPLICANT: Tumwasorn, Somying
; APPLICANT: Lepine, Guylaine
; APPLICANT: Han, Naiming
; APPLICANT: Lantz, Marilyn
; APPLICANT: Patli, Joseph
; TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
; TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ted W. Whitlock
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/570,311
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/353,485
; FILING DATE: 09-DEC-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/647,119
; FILING DATE: 25-JAN-1991
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/241,640

;; FILING DATE: 08-SEP-1988
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Whitlock, Ted W.
;; REGISTRATION NUMBER: 36,965
;; REFERENCE/DOCKET NUMBER: UF15.C3
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (904) 375-8100
;; TELEFAX: (904) 372-5800
;; INFORMATION FOR SEQ ID NO: 14:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2628 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-570-311-14

Query Match 100.0%; Score 78; DB 2; Length 2628;
Best Local Similarity 100.0%; Pred. No. 5.5e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ALNPDNYLSKDVGT 15
Db 744 ALNPDNYLSKDVGT 758

RESULT 15
US-08-570-311-2
; Sequence 2, Application US/08570311
; Patent No. 5824791
; GENERAL INFORMATION:
; APPLICANT: Proguiske-Fox, Ann
; APPLICANT: Tumwasorn, Somying
; APPLICANT: Lepine, Guylaine
; APPLICANT: Han, Naiming
; APPLICANT: Lantz, Marilyn
; APPLICANT: Patli, Joseph
; TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
; TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ted W. Whitlock
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/570,311
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/353,485
; FILING DATE: 09-DEC-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/647,119
; FILING DATE: 25-JAN-1991
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/241,640
; FILING DATE: 08-SEP-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Whitlock, Ted W.
; REGISTRATION NUMBER: 36,965
; REFERENCE/DOCKET NUMBER: UF15.C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (904) 375-8100
; TELEFAX: (904) 372-5800

; INFORMATION FOR SEQ ID NO: 2;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 497 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-570-311-2

Query Match 59.0%; Score 46; DB 2; Length 497;
Best Local Similarity 61.5%; Pred. No. 3.5;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 LNPDNVLSKDYR 14
|:|||||::||
Db 291 LSPDNVLTTPKVT 303

Search completed: June 26, 2003, 00:53:58
Job time : 2.20568 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 25, 2003, 23:05:39 ; Search time 181 Seconds

(without alignments)
591.898 Million cell updates/sec

Title: US-09-980-370-5

Perfect score: 731
Sequence: 1 gcagactccacggaacglt.....ccacggaacgaatgcgccc 402

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 1816940

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters: -DEV=xlh
-MODEL=frame_n2p.model -US09980370/runat.19062003.174442.8837/app_query.fasta_1.583
-Q/cgn2.1/USPTO.spool/US09980370 -SUFFIX=rag -MINMATCH=0.1 -DOPEL=0
-DB-A.Geneseq.101002 -OPMT=fastan -SUFFIX=rag -MINMATCH=0.1 -DOPEL=0
-LOOPEXT=0 -NITS=hits -STAR=1 -END=1 -MATRIX=Diosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pro -NOR=ext -HEAPSIE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09980370_ECGN.1.1.53_@runat.19062003.174442.8837 -NCP=6 -ICPU=3
-NO_MAP -LANG=QUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCGAPOP=6
-FCGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

A.Geneseq.101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	726	99.3	134	22	peptide used in th
2	716	97.9	1687	17	P. gingivalis haem
3	716	97.9	1687	19	Haemagglutinin pro
4	716	97.9	1704	16	Arg-gingipain-2 pr
5	716	97.9	1704	18	Arg-gingipain high
6	716	97.9	1704	21	Arg-gingipain-2 am
7	716	97.9	1704	22	P. gingivalis high
8	715	97.8	1732	17	P. gingivalis porp
9	715	97.8	1732	18	prtk antigenic pro
10	715	97.8	1732	19	Haemagglutinin pro
11	714	97.7	1706	18	P. gingivalis haem
12	711	97.3	439	17	Prtr antigenic pro
13	711	97.3	439	19	P. gingivalis haem
14	711	97.3	450	17	Haemagglutinin pro
15	711	97.3	450	19	P. gingivalis haem
16	711	97.3	450	19	Haemagglutinin pro
17	711	97.3	2628	17	P. gingivalis haem
18	709	97.0	1087	17	Haemagglutinin pro
19	709	97.0	1087	19	P. gingivalis haem
20	709	97.0	1358	17	Haemagglutinin pro
21	709	97.0	1358	19	P. gingivalis haem
22	704	96.3	456	17	Haemagglutinin pro
23	704	96.3	456	17	P. gingivalis haem
24	704	96.3	456	19	Porphyromonas ging
25	704	96.3	456	19	Haemagglutinin pro
26	205	28.0	921	20	Porphyromonas ging
27	205	28.0	922	20	Porphyromonas ging
28	205	28.0	925	20	Porphyromonas ging
29	205	28.0	938	20	Porphyromonas ging
30	180	24.6	312	20	Porphyromonas ging
31	180	24.6	312	20	Porphyromonas ging
32	176.5	24.1	497	17	Haemagglutinin pro
33	176.5	24.1	497	17	Haemagglutinin pro
34	175	23.9	419	22	Haemagglutinin pro
35	172.5	23.6	419	22	P. gingivalis haem
36	129	17.6	24	18	P. gingivalis reco
37	113	15.5	20	22	Adhesins Ptk15 an
38	108	14.8	21	22	Peptide #2. Unde
39	95.5	13.1	196	22	Peptide #3. Unde
40	94	12.9	16	20	P. gingivalis chim
41	90.5	12.4	286	22	Peptide from the 3
42	86.5	11.8	552	12	Heat-resistant neu
43	84.5	11.6	316	17	B. steatothermophil
44	83.5	11.4	316	15	Ser 144 mutant the
45	83.5	11.4	316	16	Thermolysin. Bac

ALIGNMENTS

RESULT 1
AAB49217
ID AAB49217 standard; protein: 134 AA.
XX
AC AAB49217;
XX
DT 13-MAR-2001 (first entry)
XX
DE Peptide used in the invention.
XX
KW HA2 domain; porphyrin; periodontal; pulmonary; vaginal; urethral;
infection.
XX
OS Unidentified.
XX
PN W0200072875-A1.
XX
PD 07-DEC-2000.
XX

OY 361 GCGGAGCAAGATTCGCTTTCACGAGCAATGGCC 402
 DB 1240 GYGLYAlAlArpNeglyLeuSerThrGluAlaSnGlyAla 1253

RESULT 3

AAW69495 standard; Protein: 1687 AA.
 ID AAW69495

AC AAW69495;

XX 22-DEC-1998 (first entry)

DE Haemagglutinin protein haeg.

XX Haemagglutinin protein; periodontal disease; vaccine; haeg.

XX Porphyromonas gingivalis.

XX US5824791-A.

XX 20-OCT-1998.

XX 11-DEC-1995; 95US-0570311.

XX 11-DEC-1995; 95US-0570311.

XX 08-SEP-1988; 88US-0241640.

XX 25-JAN-1991; 91US-0647119.

XX 09-DEC-1994; 94US-0353485.

XX (UABR-) UAB RES FOUND.

XX (UYFL) UNIV FLORIDA.

XX Han N. Lantz M, Lepine G, Patel JM, Progulsk-Fox A.

XX Tumescor S;

XX WPI: 1998-582627/49.

XX N-PSDB: AAV58881.

XX Isolated Porphyromonas gingivalis genes - encoding haemagglutinin

XX and/or protease poly(peptide(s))

XX Claim 1: Column 167-182; 101pp: English.

XX This sequence is encoded by a Porphyromonas gingivalis gene of the

XX invention. This sequence represents the haeg haemagglutinin protein. The

XX polypeptides are used to produce antibodies to organisms associated with

XX periodontal disease. The antibodies are also used in purification and

XX identification procedures. The genes and polypeptides are used as

XX vaccines against periodontal disease.

XX Sequence 1687 AA;

XX Alignment Scores:

XX Pred. No.: 2,91e-77

XX Score: 716.00

XX Percent Similarity: 99.25%

XX Best Local Similarity: 97.76%

XX Query Match: 97.95%

XX Db: 19

XX Gaps: 0

US-09-980-370-5 (1-402) x AAW69495 (1-1687)

OY 1 GCAGACTTCAGCAAAACCTTCAGCTTCTACATGAGAGGACACGAGGATGACT 60
 DB 1120 AAlaAspPheThrGluThrPheGluSerSerThrIleGlyAlaProAlaGluTrpThr 1139
 OY 61 ACTATGATGGCGGATGGATGGATGGATGGATGGATGGATGGATGGATGGATGGAT 120
 DB 1140 ThrIleAspPheAlaAspGlyAspGlyGlnGlyTrpLeuGlyLeuSerSerGlyGlnLeuAsp 1159
 OY 121 TGGCTACAGCTCATGCGGCGGACCAACAGTAAGTCTTCTTCATGAGATGAGTGCCT 180

DB 1160 TrpLeuThrAlaHisGlyGlyThrAsnValValAlaSerPheSerThrPasnGlyMetAla 1179

OY 181 TTGAATCCTGATACATATCTCATCTCAAGAGATGTTACAGCGGCAACGAGTAAGTAC 240

DB 1180 LeuAsnProAspAspTrpLeuIleSerLysAspValThrIleAlaThrLysValLysThr 1199

OY 241 TACTATCCAGTCACAGCAGGTTTCCCGGATCAGTACCGGTGATGATCTCAAGACG 300

DB 1200 TyrTrpAlaValAlaAspGlyPheProGlyAspHisTrpAlaValMetLleSerLysThr 1219

OY 301 GCGAGCAAGCGCGGAGCTTCAGCTTCTTCAGCAAAACGCTCAACGATTAATAG 360

DB 1220 GlyThrAsnAlaGlyAspPheThrValValPheGluGluThrProAsnGlyLleAsnLys 1239

OY 361 GCGGAGCAAGATTCGCTTTCACGAGCAATGGCC 402

DB 1240 GYGLYAlAlArpNeglyLeuSerThrGluAlaSnGlyAla 1253

RESULT 4

AA70188 standard; Protein: 1704 AA.

ID AA70188;

XX 21-SEP-1995 (first entry)

XX Arg-gingipain-2 prepolypeptide.

XX Arg-gingipain-2; gingivalis; periodontal disease; vaccine;

XX arginine-specific protease.

XX Porphyromonas gingivalis.

XX Key location/Qualifiers

XX Protein 228..719

XX FT /label="Protease

XX FT /note="corresponds to Arg-gingipain-1"

XX FT Region 720..1091

XX FT /label= Hemagglutinin

XX FT Region 1092..1429

XX FT /label= Hemagglutinin

XX FT Region 1430..1704

XX FT /label= Hemagglutinin

XX W09507286-A.

XX 16-MAR-1995.

XX 09-SEP-1994; 94MO-US10283.

XX 24-JUN-1994; 94US-0265441.

XX 10-SEP-1993; 93US-0119361.

XX 21-OCT-1993; 93US-0141324.

XX (UYGE-) UNIV GEORGIA RES FOUND INC.

XX Barr PJ, Pavloff N, Potempa J, Travis J;

XX WPI: 1995-123373/16.

XX N-PSDB: AAO83489.

XX DNA encoding Arg-gingipain proteins - used to develop probes for

XX detection, treatment and prevention of periodontal disease

XX Disclosure: Page 70-77; 89pp: English.

XX A low mol. wt. arginine-specific gingipain (AG-1) and high mol. wt.

XX AG (AG-2) were isolated from P. gingivalis strains H66 (ATCC

XX 33277) and H50 (ATCC 53973). The sequences of the proteins were

XX used to design PCR primers and probes to isolate AG DNA. Lambda

XX DASH and Lambda Zap libraries were screened with a probe based on

XX amino acids 11-22 of the AG protein to obtain DNA encoding AG-1

XX (AAO83484) and AG-2 (AAO83489). AG-2 is a prepolyprotein

CC Incorporating AG-1.
 XX Sequence 1704 AA:

Alignment Scores:

Pred. No.: 2,92e-77 Length: 1704
 Score: 716.00 Matches: 131
 Percent Similarity: 99.25% Conservative: 2
 Best Local Similarity: 97.76% Mismatches: 1
 Query Match: 97.95% Indels: 0
 DB: 16 Gaps: 0

US-09-980-370-5 (1-402) x AAK70188 (1-1704)

OY 1 GCAGACTTCACGGAAGAGCTTCATCTGATGAGAGAGCCAGCGAATGGACT 60
 DB 1137 ALaasppherhrgturhphglsuserserthrlsclglnlaaProlaoglturptr 1156
 OY 61 ACTATCGATCCGATGCGCATGCTGAGGTTGGCTGTCTGTCTTCGGAATGGAC 120
 DB 1157 ThrllaspalaaspgllyaspglnglYtrpencysleusersergylglnleuasp 1176
 OY 121 TGGCTCAGCTCTATGCGCGGACCAAGCTAGTAACTCTTCATGATGGAATGGCT 180
 DB 1177 TtrpentrallahlsiglygltthrasnvalaAlaserphesertrpansnglmetala 1196
 OY 181 TTGAATCCTGATTAACATCTCATCTCAAGAGATGTTACAGCGCAAGGTAAGTAC 240
 DB 1197 Leuasnproaspasntyrleuilleserlysaspvalthrclylaatrhllystlyr 1216
 OY 241 TACTATCCAGTCACGAGAGGTTTCCCGGATCATATCGGATGATGATCCAGAGC 300
 DB 1217 Tyrtyrlalavalasnaaspgllypheprogllyasphlstrylalvalmetlleserlysthr 1236
 OY 301 GGCAGCAAGCCCGAGACTTCACGCTGTTTTCAGAAAGCGCTTAACGAATAATAG 360
 DB 1237 GlythrlnasnaAlaglyasphehrvalvalphedlnulthrproasnnglylleasnlys 1256
 OY 361 GCGGAGCAAGATTTCGCTCTTCCACGGAAGCCCAATGGCGCC 402
 DB 1257 Glyglyalalacrgphegylleuserthrclunlaasnnglyla 1270

RESULT 5
 AAM34843
 ID AAM34843 standard: protein: 1704 AA.
 AC AAM34843:
 DT 03-JUN-1998 (first entry)
 DE Arg-ginglpain high molecular weight prepolyprotein sequence.
 XX
 XX Arg-specific ginglpain protease; ginglivalis; periodontal disease;
 KM vaccine; infection.
 XX
 OS Porphyromonas ginglivalis.
 XX
 FT Key Location/Qualifiers
 FT Protein 1..227
 FT /note= "precursor protein"
 XX
 XX MO9734629-A1.
 XX PD 25-SEP-1992.
 XX PF 21-MAR-1997: 97WO-US04635.
 XX PR 22-MAR-1996: 96US-0013945.
 XX
 XX (MORE-) MOREHOUSE SCHOOL MEDICINE.
 XX PA (UYGE-) UNIV GEORGIA RES FOUND INC.
 XX
 XX Genco CA, Potempa J, Travis J, Genco C;

XX WPI: 1997-479993/44.
 DR N-PSDB: AAT93872.
 XX

PT Porphyromonas ginglivalis Arg-specific ginglpain protease peptide(s)
 PT - useful for protecting animals and humans from ginglivalis and
 PT periodontal diseases

Disclousure: Pages 68-73: 95pp: English.

The present sequence represents an arginine-specific protease of
 CC Porphyromonas ginglivalis. The following peptides, derived from Arg-
 CC and Lys-specific high molecular weight proteases, offer protection
 CC against infection: YTYTYVRDQ IKEGITATPE DDGATGNGHE KYEKYTAGS VSPKVC
 CC (I); YYPEVKONG RMIVYAKRY (II); QLEPFDVAC VNGDFLESMR CFBAELKRAO
 CC (III); GEPNPYQPS NLRTATGQOK VTLKMDAPSTR (IV); GMEYCEVYK YTAGVSPKVC
 CC KDVTY (V); RMEFMYEPGR YYPEVKONG (VI); TFAGFEDYK RMEFMYEPGR (VII);
 CC DYTYTYVRDQ TKIKELTAT TFEEDGATG NMEYCVAKRY TAGVSPKVC (VIII); YTYTYVRDGT
 CC KIKEGLATPE EEDG (IX); RQGTIKELG TATFEEDGV ATGN (X); KIKEGLATAT
 CC FEEDGATGN HEY (XI); KMDAPNGTEN PNP NPNPN PGTTLSE (XII); and
 CC YYPEVKONG RMIVYAKRY (XIII). They are used in vaccines to protect
 CC animals, including humans, from ginglivalis and/or periodontal
 CC diseases.

XX Sequence 1704 AA:

Alignment Scores:

Pred. No.: 2,92e-77 Length: 1704
 Score: 716.00 Matches: 131
 Percent Similarity: 99.25% Conservative: 2
 Best Local Similarity: 97.76% Mismatches: 1
 Query Match: 97.95% Indels: 0
 DB: 18 Gaps: 0

US-09-980-370-5 (1-402) x AAM34843 (1-1704)

OY 1 GCAGACTTCACGGAAGAGCTTCATCTGATGAGAGAGCCAGCGAATGGACT 60
 DB 1137 ALaasppherhrgturhphglsuserserthrlsclglnlaaProlaoglturptr 1156
 OY 61 ACTATCGATCCGATGCGCATGCTGAGGTTGGCTGTCTGTCTTCGGAATGGAC 120
 DB 1157 ThrllaspalaaspgllyaspglnglYtrpencysleusersergylglnleuasp 1176
 OY 121 TGGCTCAGCTCTATGCGCGGACCAAGCTAGTAACTCTTCATGATGGAATGGCT 180
 DB 1177 TtrpentrallahlsiglygltthrasnvalaAlaserphesertrpansnglmetala 1196
 OY 181 TTGAATCCTGATTAACATCTCATCTCAAGAGATGTTACAGCGCAAGGTAAGTAC 240
 DB 1197 Leuasnproaspasntyrleuilleserlysaspvalthrclylaatrhllystlyr 1216
 OY 241 TACTATCCAGTCACGAGAGGTTTCCCGGATCATATCGGATGATGATCCAGAGC 300
 DB 1217 Tyrtyrlalavalasnaaspgllypheprogllyasphlstrylalvalmetlleserlysthr 1236
 OY 301 GGCAGCAAGCCCGAGACTTCACGCTGTTTTCAGAAAGCGCTTAACGAATAATAG 360
 DB 1237 GlythrlnasnaAlaglyasphehrvalvalphedlnulthrproasnnglylleasnlys 1256
 OY 361 GCGGAGCAAGATTTCGCTCTTCCACGGAAGCCCAATGGCGCC 402
 DB 1257 Glyglyalalacrgphegylleuserthrclunlaasnnglyla 1270

RESULT 6
 AAT67396
 ID AAT67396 standard: protein: 1704 AA.
 AC AAT67396:
 DT 25-APR-2000 (first entry)
 DE Arg-ginglpain-2 amino acid sequence.

Oy	241	AACTACTCATTCACAGACGGTTTTCCCGGGGATCACTATCGGTGATGATCTCAGACG	300
Oy	1217	TyTtYrLmAlaValaAsnSpolYpHeProGLyAphRISyTyrAlaValaMetIleSerLysThr	1236
Oy	301	GGCGAGAACCGCGGAGACTCATCGGTTTTCGAAAGAACGCCCAAGCGAATTAATAG	360
Db	1237	GlythIRmSnlaGLyAspHeThValaYlPheGLuGLuThrProAaNGlyIleAsnLys	1256
Oy	361	GGCGAGACAAGATTCGCTTTTCCACGGAACCAATAGCGCC	402
Db	1257	GlySYlAlaArpHeGLyLueSerThrGLuAlaAsnGLyAla	1270
RESULT 7			
AAU08938	AAU08938	standard; protein; 1704 AA.	
AC	AAU08938;		
AC	18-DEC-2001	(first entry)	
DE	P.	gingivalis high molecular weight Arg-gingipain-2.	
XX	XX	Periodontitis; antiinflammatory; Arg-gingipain-2; Ag-2; immunogen.	
XX	XX	Porphyromonas gingivalis.	
OS	XX		
XX	Key	Location/Qualifiers	
FH	Peptide	1..227	
FT	Region	/label= "Prepro-peptide	
FT	Region	670..674	
FT	Cleavage-site	/label= "Proteolytic-component	
FT	Protein	719	
FT	Protein	720..1091	
FT	Region	/label= "HGP_44kDa	
FT	Region	/note= "Haemagglutinin protein component"	
FT	Region	599..619	
FT	Cleavage-site	/note= "Region of homology with cysteine proteases"	
FT	Protein	1091	
FT	Protein	1092..1429	
FT	Protein	/label= "HGP_17kDa	
FT	Cleavage-site	/note= "Haemagglutinin protein component"	
FT	Protein	1429	
FT	Protein	1430..1704	
FT	Protein	/label= "HGP_27kDa	
FT	Protein	/note= "Haemagglutinin protein component"	
XX	XX		
XX	XX	US6274718-B1.	
XX	XX		
XX	XX	14-AUG-2001.	
XX	XX		
XX	XX	25-JAN-2000; 2000US-0490931.	
XX	XX		
XX	XX	24-JUN-1994; 94US-0265441.	
XX	XX	08-NOV-1994; 94US-0336308.	
XX	XX	10-SEP-1993; 93US-0119361.	
XX	XX	09-SEP-1994; 94WO-US10283.	
XX	XX		
XX	XX	(UYGE-) UNIV GEORGIA RES FOUND INC.	
XX	XX		
XX	XX	Travis J, Potempa JS, Barr PJ, Pavloff N;	
XX	XX	WPI: 2001-588904/66.	
XX	XX	DR N-PSDB: AAS15242.	
XX	XX		
XX	XX	New recombinant DNA molecule which encodes high molecular weight	
XX	XX	(mature) Arg-gingipain protein, useful for immunisation against	
XX	XX	inflammation and tissue damage, comprises enzymatically active protease	
XX	XX	component and haemagglutinin component	
XX	XX		
XX	XX	Claim 1: Column 29-41; 56pp: English.	
XX	XX		
CC	XX	The invention relates to a recombinant DNA molecule encoding high	

CC molecular weight (mature) Arg-gingipain (AG) protein, which has an
 CC enzymatically active protease component (AG-2) and a haemagglutinin
 CC component, from P. gingivalis. The nucleic acid is useful for producing
 CC mature Arg-gingipain protein. Immunogenic compositions comprising
 CC Arg-gingipain are useful for immunising animals including humans against
 CC inflammatory response and tissue damage caused by an archaebacterium
 CC Porphyromonas gingivalis, which causes progressive periodontitis.
 CC Arg-gingipain is also useful for identifying agents that modulate
 CC itself or preventing the interaction of the protease with the protein
 CC in the gingival area, such as complement factors C3 or C5. The
 CC present sequence is Arg-gingipain-2.

XX Sequence 1704 AA:

SO

Alignment Scores:

Pred. No.:	Length:	1704
Score:	716.00	131
Percent Similarity:	99.25%	2
Best Local Similarity:	97.76%	1
Query Match:	97.95%	0
DB:	22	0

US-09-980-370-5 (1-402) x AAU08938 (1-1704)

OY 1 GCAGACTTCACGGAAGATTGCGAGTCTTCTACATGAGAGGACGACGGAATGCACT 60
 |||
 DB 1137 AlaaSpheRhrGluRhrPheGluSerSerThrHisGlyAlaProAlaGluTrpThr 1156
 OY 61 ACTATCGATGCGCCATGGCCATGGTGGAGTGGCTGCTGCTGCTTCCGACATTTGGAC 120
 |||
 DB 1157 ThrileSpAlaSpolyspGlyGlnGlyTrpLeuGlySerLeuSerGlyGlnLeuasp 1176
 OY 121 TGGCTCACAGCTCACTGGCGGACCAACGTAAGTCTTCTCATGATGGAATGGCT 180
 |||
 DB 1177 TrpleuThrAlaHisGlyGlyThrAsnValAlaIaSerPheSerTrpAsnGlyMetAla 1196
 OY 181 TTGAATCCGATTAATCTCATCTCAATCAAGATGTTACAGCGCAAGAGTAATGATAC 240
 |||
 DB 1197 LeuAsnProAspAsnTrpLeuLeuSerLysAspValThrGlyAlaThrLysValLysTyr 1216
 OY 241 TACTATCCAGTCAACGACGCTTTCCGCGGATCACTAGTCCGATGATCTCCAAAGC 300
 |||
 DB 1217 TyrTrpAlaValAsnAspGlyPheProGlyAspHisTyrAlaValMetLysSerLysThr 1236
 OY 301 GCGACGAACGCGGAGACTTCACGCGTGTTCGAAAGAGCGCTTAAGCAATTAATAG 360
 |||
 DB 1237 GlyThrAsnAlaGlyAspPheThrValAlaPheGluGluTrpProAsnGlyLysLys 1256
 OY 361 GCGGAGGAGCAAGATTGCGTCTTCCAGGAGCAAGCAATGCGCC 402
 |||
 DB 1257 GlyGlyAlaArgPheGlyLeuSerThrGluAlaAsnGlyAla 1270

RESULT 8
 AAR96029
 ID AAR96029 standard; Protein: 1732 AA.
 XX
 AC AAR96029:
 XX
 DT 04-SEP-1996 (first entry)
 XX
 DE P. gingivalis porphylin.
 XX
 KW Porphylin; haemagglutinin; periodontal disease; vaccine; antibody.
 XX
 OS Porphyromonas gingivalis strain W12.
 XX
 FH Key Location/Qualifiers
 FT Region 688..708
 FT /note="Pro-Asn repeat region type 1"
 FT Region 887..952
 FT /note="Pro-Asn repeat region type 2"
 FT Region 946..967

FT
 FT Region
 FT /note="Pro-Asn repeat region type 1"
 FT 985..1006
 FT /note="Pro-Asn repeat region type 3"
 FT 1041..1100
 FT Region
 FT /note="Pro-Asn repeat region type 4"
 FT 1341..1405
 FT Region
 FT /note="Pro-Asn repeat region type 2"
 FT 1430..1451
 FT Region
 FT /note="Pro-Asn repeat region type 3"
 FT 1488..1547
 FT Region
 FT /note="Pro-Asn repeat region type 4"
 FT 1607..1650
 FT /note="Pro-Asn repeat region type 2"

XX
 XX WO9617936-A2.
 XX
 XX 13-JUN-1996.
 XX
 XX 11-DEC-1995; 95WO-US16108.
 XX
 XX 09-DEC-1994; 94US-0353485.
 XX
 XX (UABR-) UAB RES. FOUND.
 XX (UYFL) UNIV FLORIDA.
 XX
 XX Han N. Lantz M, Lepine G, Patti JM, Prognuske-Fox A;
 XX Tmwasorn S;
 XX WPI: 1996-287181/29.
 XX N-PDB: AAT30653.
 XX
 XX Porphyromonas gingivalis genes and proteins - used in the detection
 XX and vaccination against periodontal disease
 XX
 XX Claim 5: Page 76-81: 153pp: English.

CC P. gingivalis W12 cysteine protease, porphylin (AAR96029), was
 CC identified as the product of the prt gene (AAT30653) isolated from
 CC P. gingivalis W12 genomic DNA. The porphylin shows homology to
 CC the haemagglutinins (see also AAR96026-28 and AAR96030-33) of P.
 CC gingivalis 318. It can be obtained from transformed host cells and
 CC used as a vaccine to protect humans or animals against periodontal
 CC disease. Expression in salmonella cells allows production of a live
 CC vaccine. The porphylin and haemagglutinins can also be used to
 CC detect the presence of anti-P. gingivalis antibodies and to raise
 CC monoclonal antibodies for diagnostic applications.

XX
 SO Sequence 1732 AA:

Alignment Scores:

Pred. No.:	Length:	1732
Score:	715.00	131
Percent Similarity:	99.25%	2
Best Local Similarity:	97.76%	1
Query Match:	97.81%	0
DB:	17	0

US-09-980-370-5 (1-402) x AAR96029 (1-1732)

OY 1 GCAGACTTCACGGAAGATTGCGAGTCTTCTACATGAGAGGACGACGGAATGCACT 60
 |||
 DB 1157 AlaaSpheRhrGluRhrPheGluSerSerThrHisGlyAlaProAlaGluTrpThr 1176
 OY 61 ACTATCGATGCGCCATGGCCATGGTGGAGTGGCTGCTGCTGCTTCCGACATTTGGAC 120
 |||
 DB 1177 ThrileSpAlaSpolyspGlyGlnGlyTrpLeuGlySerLeuSerGlyGlnLeuasp 1196
 OY 121 TGGCTCACAGCTCACTGGCGGACCAACGTAAGTCTTCTCATGATGGAATGGCT 180
 |||
 DB 1197 TrpleuThrAlaHisGlyGlyThrAsnValAlaIaSerPheSerTrpAsnGlyMetAla 1216
 OY 181 TTGAATCCGATTAATCTCATCTCAATCAAGATGTTACAGCGCAAGAGTAATGATAC 240
 |||

CC other previously prepared antigens based on fimbriae or the
CC capsule, the Pritr-PrtK complex or component parts are safe and
CC effective antigens.

SO Sequence 1706 AA:

Alignment Scores:
Pred. No.: 5,09e-77 Length: 1706
Score: 714.00 Matches: 131
Percent Similarity: 99.25% Conservative: 2
Best Local Similarity: 97.76% Mismatches: 1
Query Match: 97.67% Indels: 0
DB: 18 Gaps: 0

US-09-980-370-5 (1-402) x AAM24786 (1-1706)

OY 1 GCAGACTTCACGGAAGCTTCGAGTCTCTACATGAGAGGACCCAGCGAATGCACT 60
DB 1139 AAlasprhethrGluThrPheGluSerSerThrHisGlyAlaProAlaGluTrpThr 1158
OY 61 ACTATCGATCCGATGCGGATGATGAGGCTGCGCTCTGTCGCGACATTTGGAC 120
DB 1159 ThrIleAspAlaAspGlyAspGlyGlnGlyTrpLeuGlySerSerGlyGlnLeuAsp 1178
OY 121 TGCGTCACAGCTCATGCGGCGACACAGCTAGTAAGCTCTTCTCATGGAATGAGCT 180
DB 1179 TrpLeuThrAlaHisGlyGlyThrAsnValAlaSerSerPheSerTrpAsnGlyMetAla 1198
OY 181 TTGAATCCTGATTAATCTATCTCATCAAGAGATGTTACAGCGGACAGAGTAAGTAC 240
DB 1199 LeuAsnProAspAsnTyrlleuIleSerIysAspValThrGlyAlaThrIysValIysTyr 1218
OY 241 TACTATCCAGTCAACGAGGTTTCCGCGGATCACTATGCGGATGATCTCCAGACG 300
DB 1219 TyrlleAlaValAlaAspGlyPheProGlyAspHisTyrAlaValMetIleSerIysThr 1238
OY 301 GGCACGACGCGGACAGCTTACGCGTGTCTTTCAGAGAACGCGCTAACGGAATTAATAG 360
DB 1239 GlyThrAsnAlaGlyAspPheThrValAlaPheGluGluThrProAsnGlyIleAsnIys 1258
OY 361 GCGGAGCAGAGATTCGCTCTTCCAGGAGAACCCAAATGCGCGC 402
DB 1259 GlyGlyAlaAspPheGlyLeuSerThrGluAlaAspGlyAla 1272
RESULT 12
ID AAR96024 standard; Protein: 439 AA.
XX AAR96024:
XX 04-SEP-1996 (first entry)
XX P. gingivalis haemagglutinin haga Harep4 product. *d.f.*
XX DE Haemagglutinin; haga: periodontal disease; vaccine; antibody;
XX KW Harep4.
XX OS Porphyromonas gingivalis strain 381.
XX PN MO9617936-A2.
XX PD 13-JUN-1996.
XX 11-DEC-1995: 95WO-US16108.
XX PR 09-DEC-1994: 94US-0353485.
XX (UABR-) UAB RES FOUND.
XX (UFL) UNIV FLORIDA.
XX Han N, Lantz M, Lepine G, Patel JM, Progulsk-Fox A;
XX Tumwaborn S;

DR WPI: 1996-287181/29;
DR N-PSDB: AAT30648.

PT Porphyromonas gingivalis genes and proteins - used in the detection
PT and vaccination against periodontal disease

PS Claim 4; Page 114-115; 153pp; English.

XX Harep4 (AAR96024) is the product of the Harep4 repeat unit (AAT30648)
CC of the haga gene (AAT30648) of P. gingivalis 318. It forms part
CC of haemagglutinin haga (see also AAR96030). Harep4 and other
CC haga repeat unit products (see also AAR96021-23) can be obtd. from
CC transformed host cells and used as vaccines to protect humans or
CC animals against periodontal disease. Expression in Salmonella
CC cells allows prodn. of live vaccine. Harep4 can also be used
CC to detect the presence of anti-P. gingivalis antibodies and to
CC raise monoclonal antibodies for diagnostic appln.

SO Sequence 439 AA:

Alignment Scores:
Pred. No.: 7.15e-77 Length: 439
Score: 711.00 Matches: 130
Percent Similarity: 99.25% Conservative: 3
Best Local Similarity: 97.01% Mismatches: 1
Query Match: 97.26% Indels: 0
DB: 17 Gaps: 0

US-09-980-370-5 (1-402) x AAR96024 (1-439)

OY 1 GCAGACTTCACGGAAGCTTCGAGTCTCTACATGAGAGGACCCAGCGAATGCACT 60
DB 192 AAlasprhethrGluThrPheGluSerSerThrHisGlyAlaProAlaGluTrpThr 211
OY 61 ACTATCGATCCGATGCGGATGATGAGGCTGCGCTCTGTCGCGACATTTGGAC 120
DB 212 ThrIleAspAlaAspGlyAspGlyGlnGlyTrpLeuGlySerSerGlyGlnLeuAsp 231
OY 121 TGCGTCACAGCTCATGCGGCGACACAGCTAGTAAGCTCTTCTCATGGAATGAGCT 180
DB 232 TrpLeuThrAlaHisGlyGlyThrAsnValAlaSerSerPheSerTrpAsnGlyMetAla 251
OY 181 TTGAATCCTGATTAATCTATCTCATCAAGAGATGTTACAGCGGACAGAGTAAGTAC 240
DB 252 LeuAsnProAspAsnTyrlleuIleSerIysAspValThrGlyAlaThrIysValIysTyr 271
OY 241 TACTATCCAGTCAACGAGGTTTCCGCGGATCACTATGCGGATGATCTCCAGACG 300
DB 272 TyrlleAlaValAlaAspGlyPheProGlyAspHisTyrAlaValMetIleSerIysThr 291
OY 301 GGCACGACGCGGACAGCTTACGCGTGTCTTTCAGAGAACGCGCTAACGGAATTAATAG 360
DB 292 GlyThrAsnAlaGlyAspPheThrValAlaPheGluGluThrProAsnGlyIleAsnIys 311
OY 361 GCGGAGCAGAGATTCGCTCTTCCAGGAGAACCCAAATGCGCGC 402
DB 312 GlyGlyAlaAspPheGlyLeuSerThrGluAlaAspGlyAla 325
RESULT 13
ID AAM69492 standard; Protein: 439 AA.
XX AAM69492:
XX 22-DEC-1998 (first entry)
XX Haemagglutinin protein haga, Harep4.
XX DE Haemagglutinin protein; haga: periodontal disease; vaccine; haga.
XX KW Haemagglutinin protein; haga: periodontal disease; vaccine; haga.
XX OS Porphyromonas gingivalis.
XX PN US5824791-A.

XX 20-OCT-1998.
PD
XX
XX 11-DEC-1995; 95US-0570311.
PF
XX 11-DEC-1995; 95US-0570311.
PR
XX 08-SEP-1988; 88US-0241640.
PR 25-JAN-1991; 91US-0647119.
PR 09-DEC-1994; 94US-0353485.
XX
XX (UABR-) UAB RES FOUND.
PA (UYFL) UNITV FLORIDA.
XX
XX Han N, Lantiz M, Lepine G, Patti JM, Progulske-Fox A;
PI Tumasorn S;
XX
XX WPI: 1998-582627/49.
DR N-PSDB; AAV58879.
XX
XX Isolated Porphyromonas gingivalis genes - encoding haemagglutinin
PT and/or protease poly(peptide(s))
XX
XX Claim 1; Column 139-144; 101pp; English.
XX
XX This sequence is encoded by a Porphyromonas gingivalis gene of the
CC invention. This sequence represents the hagd haemagglutinin protein. The
CC polypeptides are used to produce antibodies to organisms associated with
CC periodontal disease. The antibodies are also used in purification and
CC identification procedures. The genes and polypeptides are used as
CC vaccines against periodontal disease.
XX
SQ Sequence 439 AA;

Alignment Scores:
Pred. No.: 7,15e-77 Length: 439
Score: 711.00 Matches: 130
Percent Similarity: 99.25% Conservative: 3
Best Local Similarity: 97.01% Mismatches: 1
Query Match: 97.26% Indels: 0
DB: 19 Gaps: 0

US-09-980-370-5 (1-402) x AAV69492 (1-439)

OY 1 GCAGACTTCACGGAACGTTGCGAGTCTTCTACTCATGAGAGGACGACCGAATGGACT 60
DB 192 AlaAspRheThrCluThrPheGluSerSerThrHisGlyAlaProIaGluTrpThr 211
OY 61 ACTATCGATGCGGATGGCGATGGTGAGGGTTGGCTGTCTGTCTCCGACAAATTGAC 120
DB 212 ThrIleAspAlaAspGlyAspGlyGlnGlyTrpLeuGlyLeuSerSerGlyGlnLeuAsp 231
OY 121 TGGGTCACAGCTCATGCGCGGACGACCAAGTAGAGCTTTTTCATGGAATGGAGCT 180
DB 222 TrpLeuThrAlaHisGlyGlyThrAsnValValAlaIleSerPheSerThrAsnGlyMetAla 251
OY 181 TTGAATCCTGATACATCTCATCTCAAGAGATGTTACGAGGCGCAAGCAAGTAAAGTAC 240
DB 252 LeuAsnProAspAsnTyrIleuIleSerIysAspValThrGlyAlaThrIlyValIlyTyr 271
OY 211 TACTATCGACATCAAGCAGCGGTTTCCCGGGATCACTATGCGGTGATGTTCCAAAGAC 300
DB 272 TgtTgtAlaValaAspGlyAspGlyAspGlyAspGlyAspGlyAspGlyAspGlyAsp 291
OY 301 GGCAGAGACCGCGGAGACTTCACGCTGTTTTCGAAGAAGCCGTAAGCGAATGAATAG 360
DB 292 GlyThrAsnAlaGlyAspRheThrValValIleGlnGlyIleThrProAsnGlyIleAsnIys 311
OY 361 GGCAGAGCAAGATGCTGTTCTTCCACGAGCAAGCAATGGCGCC 402
DB 312 GlyGlyAlaIleArgPheGlyLeuSerThrGlyAlaIleAspGlyAla 325
RESULT 14
AA96021

ID AAR96021 standard; Protein; 450 AA.
XX
XX AAR96021:
XX
XX 04-SEP-1996 (first entry)
XX
XX P. gingivalis haemagglutinin hagd HAREPI product.
DE
XX Haemagglutinin; hagd; periodontal disease; vaccine; antibody;
KW HAREPI.
XX
XX OS Porphyromonas gingivalis strain 381.
XX
XX WO9617936-A2.
XX
XX 13-JUN-1996.
XX
XX 11-DEC-1995; 95WO-US16108.
XX
XX 09-DEC-1994; 94US-0353485.
XX
XX (UABR-) UAB RES FOUND.
PA (UYFL) UNITV FLORIDA.
XX
XX Han N, Lantiz M, Lepine G, Patti JM, Progulske-Fox A;
PI Tumasorn S;
XX
XX WPI: 1996-287181/29.
DR N-PSDB; AAT30645.
XX
XX Porphyromonas gingivalis genes and proteins - used in the detection
PT and vaccination against periodontal disease
XX
XX Claim 4; Page 103-104; 153pp; English.
XX
XX HAREPI (AAR96021) is the product of the HAREPI repeat unit (AAT30645)
CC of the hagd gene (AAT30654) of P. gingivalis 318. It forms part
CC of haemagglutinin hagd (see also AAR96030). HAREPI and other
CC hagd repeat unit products (see also AAR96022-24) can be obtd. from
CC transformed host cells and used as vaccines to protect humans or
CC animals against periodontal disease. Expression in Salmonella
CC cells allows prodn. of live vaccine. HAREPI-4 can also be used
CC to detect the presence of anti-P. gingivalis antibodies and to
CC raise monoclonal antibodies for diagnostic appln.
XX
SQ Sequence 450 AA;

Alignment Scores:

Pred. No.: 7,22e-77 Length: 450
Score: 711.00 Matches: 130
Percent Similarity: 99.25% Conservative: 3
Best Local Similarity: 97.01% Mismatches: 1
Query Match: 97.26% Indels: 0
DB: 17 Gaps: 0

US-09-980-370-5 (1-402) x AAR96021 (1-450)

OY 1 GCAGACTTCACGGAACGTTGCGAGTCTTCTACTCATGAGAGGACGACCGAATGGACT 60
DB 186 AlaAspRheThrCluThrPheGluSerSerThrHisGlyAlaProIaGluTrpThr 205
OY 61 ACTATCGATGCGGATGGCGATGGTGAGGGTTGGCTGTCTGTCTCCGACAAATTGAC 120
DB 206 ThrIleAspAlaAspGlyAspGlyGlnGlyTrpLeuGlyLeuSerSerGlyGlnLeuAsp 225
OY 121 TGGGTCACAGCTCATGCGCGGACGACCAAGTAGAGCTTTTTCATGGAATGGAGCT 180
DB 226 TrpLeuThrAlaHisGlyGlyThrAsnValValAlaIleSerPheSerThrAsnGlyMetAla 245
OY 181 TTGAATCCTGATACATCTCATCTCAAGAGATGTTACGAGGCGCAAGCAAGTAAAGTAC 240
DB 246 LeuAsnProAspAsnTyrIleuIleSerIysAspValThrGlyAlaThrIlyValIlyTyr 265

OY	241	TACATTCAGTCAAGGAGCGTTTCCGGGAAATCATATGCGTGATGATCTCCAGAGC	360
Dd	266	TyrtYrAlaValaIAsnAspelyPheProGlyaspHisTyrlAlaValaMetileSerLysThr	285
OY	301	GCCACGAAGCCCGGAGACTTCACGGGTGTTTTCGAAGAAGCGCTAACGGAATAAATAAG	360
Dd	286	GlythrAsnAlaGlyAspPheThrValAlAPheGluGLIThrProAsnGlyLeasLys	305
OY	361	GGCGGAGCAGATTTCGTCTTTCACAGCAAGCCAATGGCGCC	402
Dd	306	GlyGlyAlaArgPheGlyLeuSerThrcuaIlaAspGlyAla	319

RESULT 15

ID	AAW69489 standard; protein; 450 AA.
xx	

AC AAW69489;

DT 22-DEC-1998 (first entry)

DE Haemagglutinin protein hgaA, HAREp1.

KW Haemagglutinin protein; periodontal disease; vaccine; hga

Porphyromonas gingivalis.

PN US5824791-A.

PD 20-OCT-1998.

PF 11-DEC-1995; 95US-0570311.
 XX

PR	11-DEC-1995;	95US-0570311.
DP	08-SEP-1999.	89US-0341640

PR	25-JAN-1991;	91US-0647119.
PR	09-DEC-1994;	91US-0353485.

XX
XX
/11ABP-) 11AB REC ECONUM
DA

PA (UYFL) UNIV FLORIDA.
XX

PI Han N, Lantz M, Lepine G, Patil JM, Progułske-Fox A;
PI Tumwasorn S:

XX WPT: 1000-503637/40
NB

DR N-PSDB; AAV58876.
XX

PT Isolated porphyromonas gingivalis genes - encoding haemagglutinin and/or protease polypeptides(s))

XX
PS
Claim 1: Columns 121-126; 101cc. Escalator

This sequence is encoded by a Bombus terrestris gene of the CCXX

CC Invention: This sequence represents the hga haemagglutinin protein. The CC polynucleotides are used to produce antibodies to extrinsic structural

periodontal disease. The antibodies are also used in purification and identification procedures. The cones and columns are used in

cc Vaccines against periodontal disease.
XX

sequence 450 AA;

Alignment scores:
Pred. No.:

score:	711.00
Percent similarity:	99.258

Best Local Similarity: 97.01%
Query Match: 97.26%

DB: 19

03-03-980-3/0-3 (1-402) X AAWb9489 (1-450)

1 GCGAGACACGATCGAGCTCTCTACCTCAGAGAGGCCACCAGCGGATGGACT 60

100 mdsprfneimdlrphglusersegrlphnlsdglglualafroalaglttrthr 205

[illegible]

Search completed: June 26, 2003, 00:42:49
Job time : 187 secs

Job time : 187 secs

•
1
-
•

Thu Jun 26 11:58:01 2003

us-09-980-370-1.rag

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 26, 2003, 00:44:59 ; Search time 5.29412 Seconds
(without alignments)

377.543 Million cell updates/sec

Title: US-09-980-370-1

Sequence: 1 ALNPDNYLSKDVTC 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Genseq.101002:*
1: /SID52/gcgdata/geneeq/geneeqp-emb1/AA1980.DAT:*
2: /SID52/gcgdata/geneeq/geneeqp-emb1/AA1981.DAT:*
3: /SID52/gcgdata/geneeq/geneeqp-emb1/AA1982.DAT:*
4: /SID52/gcgdata/geneeq/geneeqp-emb1/AA1983.DAT:*
5: /SID52/gcgdata/geneeq/geneeqp-emb1/AA1984.DAT:*
6: /SID52/gcgdata/geneeq/geneeqp-emb1/AA1985.DAT:*
7: /SID52/gcgdata/geneeq/geneeqp-emb1/AA1986.DAT:*
8: /SID52/gcgdata/geneeq/geneeqp-emb1/AA1987.DAT:*
9: /SID52/gcgdata/geneeq/geneeqp-emb1/AA1988.DAT:*
10: /SID52/gcgdata/geneeq/geneeqp-emb1/AA1989.DAT:*
11: /SID52/gcgdata/geneeq/geneeqp-emb1/AA1990.DAT:*
12: /SID52/gcgdata/geneeq/geneeqp-emb1/AA1991.DAT:*
13: /SID52/gcgdata/geneeq/geneeqp-emb1/AA1992.DAT:*
14: /SID52/gcgdata/geneeq/geneeqp-emb1/AA1993.DAT:*
15: /SID52/gcgdata/geneeq/geneeqp-emb1/AA1994.DAT:*
16: /SID52/gcgdata/geneeq/geneeqp-emb1/AA1995.DAT:*
17: /SID52/gcgdata/geneeq/geneeqp-emb1/AA1996.DAT:*
18: /SID52/gcgdata/geneeq/geneeqp-emb1/AA1997.DAT:*
19: /SID52/gcgdata/geneeq/geneeqp-emb1/AA1998.DAT:*
20: /SID52/gcgdata/geneeq/geneeqp-emb1/AA1999.DAT:*
21: /SID52/gcgdata/geneeq/geneeqp-emb1/AA2000.DAT:*
22: /SID52/gcgdata/geneeq/geneeqp-emb1/AA2001.DAT:*
23: /SID52/gcgdata/geneeq/geneeqp-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	78	100.0	15	AA849215	Residues 1215-1229
2	78	100.0	21	AA849218	Peptide #3, unde
3	78	100.0	134	AA849217	Peptide used in th
4	78	100.0	439	AA849024	P. gingivalis haem
5	78	100.0	439	AA849024	P. gingivalis haem
6	78	100.0	450	AA849021	P. gingivalis haem
7	78	100.0	450	AA849021	P. gingivalis haem
8	78	100.0	456	AA849022	Haemagglutinin pro
9	78	100.0	456	AA849022	P. gingivalis haem
10	78	100.0	456	AA849023	P. gingivalis haem
			19	AA849430	Haemagglutinin pro

11	78	100.0	456	19	AA849431	Haemagglutinin pro
12	78	100.0	1087	17	AA849432	P. gingivalis haem
13	78	100.0	1087	19	AA849436	Haemagglutinin pro
14	78	100.0	1358	17	AA849432	P. gingivalis haem
15	78	100.0	1358	19	AA849434	Haemagglutinin pro
16	78	100.0	1687	17	AA849433	P. gingivalis haem
17	78	100.0	1687	19	AA849435	Haemagglutinin pro
18	78	100.0	1704	16	AA849438	Haemagglutinin pro
19	78	100.0	1704	18	AA849439	Arg-gingipain-2 pr
20	78	100.0	1704	21	AA849439	Arg-gingipain-2 pr
21	78	100.0	1704	22	AA849439	Arg-gingipain-2 pr
22	78	100.0	1704	22	AA849439	Arg-gingipain-2 pr
23	78	100.0	1732	17	AA849439	P. gingivalis haem
24	78	100.0	1732	18	AA849439	P. gingivalis haem
25	78	100.0	1732	19	AA849439	P. gingivalis haem
26	78	100.0	1732	19	AA849439	P. gingivalis haem
27	78	100.0	2528	17	AA849439	Haemagglutinin pro
28	78	100.0	2528	19	AA849439	P. gingivalis haem
29	46	59.0	487	17	AA849439	Haemagglutinin pro
30	46	59.0	487	19	AA849439	Propionibacterium
31	45	57.7	159	22	AA849439	Propionibacterium
32	45	57.7	312	20	AA849439	Propionibacterium
33	45	57.7	377	20	AA849439	Propionibacterium
34	44	56.4	315	23	AA849439	Herbicideally activ
35	44	56.4	639	23	AA849439	Lactococcus lactis
36	43	55.1	518	22	AA849439	Drosophila melanog
37	43	55.1	921	20	AA849439	Propionibacterium
38	43	55.1	925	20	AA849439	Propionibacterium
39	43	55.1	938	20	AA849439	Propionibacterium
40	42	53.8	374	21	AA849439	Arabidopsis thaliana
41	42	53.8	374	21	AA849439	Arabidopsis thaliana
42	42	53.8	394	21	AA849439	Human myxoma virus
43	42	53.8	2548	21	AA849439	Human myxoma virus
44	41	52.6	69	23	AA849439	Human myxoma virus
45	41	52.6	166	21	AA849439	Human myxoma virus

ALIGNMENTS

RESULT 1
ID AA849215 standard; peptide: 15 AA.
AC AA849215;
DT 13-MAR-2001 (first entry)
DE Residues 1215-1229 of HA2 gingipain domain.
KW HA2 domain; porphyrin; periodontal; pulmonary; vaginal; urethral;
infection.
OS Porphyromonas gingivalis.
PN WO200072875-A1.
PD 07-DEC-2009.
PF 26-MAY-2000; 2000MO-AU00599.
PR 28-MAY-1999; 99AU-0000652.
XX (UNSY) UNTV SYDNEY.
XX COLLYER CA, Hunter N, De Carlo AA.
XX WPI, 2001-080424/09.
XX Treating microbial infection in environment containing porphyrin, by
XX administering a HA-2 antagonist
XX Claim 11: Page 44; 102pp; English.
PS

XX The present invention relates to prophylaxis or treatment of infection
 CC by a microorganism in a biological environment comprising iron, heme
 CC or porphyrin. The treatment involves administering to the environment,
 CC an agent which antagonizes the interaction between a molecule derived
 CC from the microorganism and having an HA2 domain, and an HA2-binding
 CC motif on a porphyrin containing molecule present in the environment.
 CC Useful in the manufacture of a medicament for the prophylaxis and
 CC treatment of periodontal, pulmonary, vaginal, urethral or hoof disease
 CC resulting from Porphyromonas gingivalis infection or infection by a
 CC related microorganism.

XX Sequence 15 AA:

Query Match 100.0%; Score 78; DB 22; Length 15;
 Best Local Similarity 100.0%; Pred. No. 1.3e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNDPNYLISKDVVG 15
 DB 1 ALNDPNYLISKDVVG 15

RESULT 2

ID AAB49218 standard; peptide: 21 AA.

AC AAB49218:

DT 13-MAR-2001 (first entry)

DE Peptide #3.

XX HA2 domain; porphyrin; periodontal; pulmonary; vaginal; urethral;
 KW infection.

OS Unidentified.

PN W0200072875-A1.

PD 07-DEC-2000.

PE 26-MAY-2000; 2000WO-AU00599.

PR 28-MAY-1999; 99AU-0000652.

FA (UNSY) UNIV SYDNEY.

XX Collyer CA, Hunter N, De Carlo AA;

WI: 2001-080424/09.

XX Treating microbial infection in environment containing porphyrin, by
 PT administering a HA-2 antagonist
 XX Claim 11: Page 40; 102pp; English.

XX The present invention relates to prophylaxis or treatment of infection
 CC by a microorganism in a biological environment comprising iron, heme
 CC or porphyrin. The treatment involves administering to the environment,
 CC an agent which antagonizes the interaction between a molecule derived
 CC from the microorganism and having an HA2 domain, and an HA2-binding
 CC motif on a porphyrin containing molecule present in the environment.
 CC Useful in the manufacture of a medicament for the prophylaxis and
 CC treatment of periodontal, pulmonary, vaginal, urethral or hoof disease
 CC resulting from Porphyromonas gingivalis infection or infection by a
 CC related microorganism.

XX Sequence 21 AA:

Query Match 100.0%; Score 78; DB 22; Length 21;
 Best Local Similarity 100.0%; Pred. No. 2e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALNDPNYLISKDVVG 15
 DB 1 ALNDPNYLISKDVVG 15

RESULT 3

ID AAB49217 standard; protein: 134 AA.

AC AAB49217:

DT 13-MAR-2001 (first entry)

DE Peptide used in the invention.

XX HA2 domain; porphyrin; periodontal; pulmonary; vaginal; urethral;
 KW infection.

OS Unidentified.

PN W0200072875-A1.

PD 07-DEC-2000.

PE 26-MAY-2000; 2000WO-AU00599.

PR 28-MAY-1999; 99AU-0000652.

FA (UNSY) UNIV SYDNEY.

XX Collyer CA, Hunter N, De Carlo AA;

WI: 2001-080424/09.

XX Treating microbial infection in environment containing porphyrin, by
 PT administering a HA-2 antagonist
 XX Claim 9: Page 98-99; 102pp; English.

XX The present invention relates to prophylaxis or treatment of infection
 CC by a microorganism in a biological environment comprising iron, heme
 CC or porphyrin. The treatment involves administering to the environment,
 CC an agent which antagonizes the interaction between a molecule derived
 CC from the microorganism and having an HA2 domain, and an HA2-binding
 CC motif on a porphyrin containing molecule present in the environment.
 CC Useful in the manufacture of a medicament for the prophylaxis and
 CC treatment of periodontal, pulmonary, vaginal, urethral or hoof disease
 CC resulting from Porphyromonas gingivalis infection or infection by a
 CC related microorganism.

XX Sequence 134 AA:

Query Match 100.0%; Score 78; DB 22; Length 134;
 Best Local Similarity 100.0%; Pred. No. 1.9e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNDPNYLISKDVVG 15
 DB 60 ALNDPNYLISKDVVG 74

RESULT 4

ID AAB96024 standard; protein: 439 AA.

AC AAB96024:

DT 04-SEP-1996 (first entry)

XX P. gingivalis haemagglutinin haga Harep4 product.

XX Haemagglutinin; haga; periodontal disease; vaccine; antibody;

KW HAREP4.
 XX Porphyromonas gingivalis strain 381.
 OS
 XX MO9617936-A2.
 PM
 XX 13-JUN-1996.
 PD
 XX 11-DEC-1995: 95WO-US16108.
 PF
 XX 09-DEC-1994: 94US-0353485.
 PR
 XX (UABR-) UAB RES FOUND.
 PA (UVFL) UNIV FLORIDA.
 XX
 PI Han N, Lantz M, Lepine G, Patti JM, Proguliske-Fox A;
 PT Tumwasorn S;
 XX WPI: 1996-287181/29.
 DR N-PSDB: AAT30648.
 DR
 XX Porphyromonas gingivalis genes and proteins - used in the detection
 PT and vaccination against periodontal disease
 PS
 XX Claim 4: Page 114-115; 153pp: English.
 CC
 XX HAREP4 (AAR96024) is the product of the HAREP4 repeat unit (AAT30648)
 CC of the hga gene (AAT30654) of P. gingivalis 318. It forms part
 CC of haemagglutinin hga (see also AAR96030). HAREP4 and other
 CC hga repeat unit products (see also AAR96021-23) can be obtd. from
 CC transformed host cells and used as vaccines to protect humans or
 CC animals against periodontal disease. Expression in Salmonella
 CC cells allows produ. of live vaccine. HAREP4 can also be used
 CC to detect the presence of anti-P. gingivalis antibodies and to
 CC raise monoclonal antibodies for diagnostic appln.
 CC
 XX
 SQ Sequence 439 AA:
 QY
 DB 1 ALNPDNYLISKDVTG 15
 251 ALNPDNYLISKDVTG 265
 RESULT 5
 AAM69492
 ID AAM69492 standard; Protein: 439 AA.
 XX
 AC AAM69492;
 XX
 DT 22-DEC-1998 (first entry)
 DE
 XX Haemagglutinin protein hga, HAREP4.
 DE
 XX Haemagglutinin protein; periodontal disease; vaccine; hga.
 KW
 XX Porphyromonas gingivalis.
 OS
 XX US5824791-A.
 PM
 XX 20-OCT-1998.
 PD
 XX 11-DEC-1995: 95US-0570311.
 PF
 XX 11-DEC-1995: 95US-0570311.
 PR 08-SEP-1988: 88US-0241640.
 PR 25-JAN-1991: 91US-0647119.
 PR 09-DEC-1994: 94US-0353485.
 XX
 PA (UABR-) UAB RES FOUND.

PA (UVFL) UNIV FLORIDA.
 XX
 PI Han N, Lantz M, Lepine G, Patti JM, Proguliske-Fox A;
 PT Tumwasorn S;
 XX WPI: 1998-582627/49.
 DR N-PSDB: AAV38879.
 DR
 XX Isolated Porphyromonas gingivalis genes - encoding haemagglutinin
 PT and/or protease polypeptide(s)
 PS
 XX Claim 1: Column 139-144; 101pp: English.
 CC
 XX This sequence is encoded by a Porphyromonas gingivalis gene of the
 CC invention. This sequence represents the hga haemagglutinin protein. The
 CC polypeptides are used to produce antibodies to organisms associated with
 CC periodontal disease. The antibodies are also used in purification and
 CC identification procedures. The genes and polypeptides are used as
 CC vaccines against periodontal disease.
 CC
 XX
 SQ Sequence 439 AA:
 QY
 DB 1 ALNPDNYLISKDVTG 15
 251 ALNPDNYLISKDVTG 265
 RESULT 6
 AAR96021
 ID AAR96021 standard; Protein: 450 AA.
 XX
 AC AAR96021;
 XX
 DT 04-SEP-1996 (first entry)
 DE
 XX P. gingivalis haemagglutinin hga HAREP1 product.
 DE
 XX Haemagglutinin: hga; periodontal disease; vaccine; antibody;
 KW HAREP1.
 XX
 XX Porphyromonas gingivalis strain 381.
 OS
 XX MO9617936-A2.
 PM
 XX 13-JUN-1996.
 PD
 XX 11-DEC-1995: 95WO-US16108.
 PF
 XX 09-DEC-1994: 94US-0353485.
 PR
 XX (UABR-) UAB RES FOUND.
 PA (UVFL) UNIV FLORIDA.
 XX
 PI Han N, Lantz M, Lepine G, Patti JM, Proguliske-Fox A;
 PT Tumwasorn S;
 XX WPI: 1996-287181/29.
 DR N-PSDB: AAT30645.
 DR
 XX Porphyromonas gingivalis genes and proteins - used in the detection
 PT and vaccination against periodontal disease
 PS
 XX Claim 4: Page 103-104; 153pp: English.
 CC
 XX HAREP1 (AAR96021) is the product of the HAREP1 repeat unit (AAT30645)
 CC of the hga gene (AAT30654) of P. gingivalis 318. It forms part
 CC of haemagglutinin hga (see also AAR96030). HAREP1 and other
 CC hga repeat unit products (see also AAR96022-24) can be obtd. from
 CC transformed host cells and used as vaccines to protect humans or

CC animals against periodontal disease. Expression in Salmonella
CC cells allows production of live vaccine. Harep1-4 can also be used
CC to detect the presence of anti-P. gingivalis antibodies and to
CC raise monoclonal antibodies for diagnostic applan.

SO Sequence 450 AA:

Query Match 100.0%: Score 78: DB 17: Length 450:
Best Local Similarity 100.0%: Pred. No. 8-2e-06:
Matches 15: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

QY 1 ALNPNVYLISKDVVG 15
DB 245 ALNPNVYLISKDVVG 259

RESULT 7

AAW69489
ID AAW69489 standard; Protein: 450 AA.

AC AAW69489:

DT 22-DEC-1998 (first entry)

DE Haemagglutinin protein haggA, Harep1.

KW Haemagglutinin protein; periodontal disease; vaccine; haggA.

OS Porphyromonas gingivalis.

PN US5824791-A.

PD 20-OCT-1998.

XX 11-DEC-1995: 95US-0570311.

PR 11-DEC-1995: 95US-0570311.

PR 08-SEP-1988: 88US-0241640.

PR 25-JAN-1991: 91US-0647119.

PR 09-DEC-1994: 94US-0353485.

PA (UABR-) UAB RES FOUND.

PI (UYFL) UNIV FLORIDA.

PI Han N. Lantz M, Lepine G, Patti JM, Proguiske-Fox A;

PI Tummasorn S;

DR WPI: 1998-582627/49.

DR N-PSDB: AAV58876.

PT Isolated Porphyromonas gingivalis genes - encoding haemagglutinin

PT and/or protease poly(peptide(s))

PS Claim 1: Column 121-126: 101pp: English.

XX This sequence is encoded by a Porphyromonas gingivalis gene of the

CC invention. This sequence represents the haggA haemagglutinin protein. The

CC polypeptides are used to produce antibodies to organisms associated with

CC periodontal disease. The antibodies are also used in purification and

CC identification procedures. The genes and polypeptides are used as

CC vaccines against periodontal disease.

SO Sequence 450 AA:

Query Match 100.0%: Score 78: DB 19: Length 450:

Best Local Similarity 100.0%: Pred. No. 8-2e-06:

Matches 15: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

QY 1 ALNPNVYLISKDVVG 15

DB 245 ALNPNVYLISKDVVG 259

RESULT 8

AAAR96022
ID AAAR96022 standard; Protein: 456 AA.

AC AAAR96022:

DT 04-SEP-1996 (first entry)

DE P. gingivalis haemagglutinin haggA Harep2 product.

KW Haemagglutinin: haggA; periodontal disease; vaccine; antibody;

KW Harep2.

OS Porphyromonas gingivalis strain 381.

PN WO9617936-A2.

PD 13-JUN-1996.

XX 11-DEC-1995: 95WO-US16108.

PR 09-DEC-1994: 94US-0353485.

PA (UABR-) UAB RES FOUND.

PI (UYFL) UNIV FLORIDA.

PI Han N. Lantz M, Lepine G, Patti JM, Proguiske-Fox A;

PI Tummasorn S;

DR WPI: 1996-287181/29.

DR N-PSDB: AAT30646.

PT Porphyromonas gingivalis genes and proteins - used in the detection

PT and vaccination against periodontal disease

PS Claim 4: Page 107-108: 153pp: English.

XX Harep2 (AAAR96022) is the product of the Harep2 repeat unit (AAT30646)

CC of the haggA gene (AAT30654) of P. gingivalis 318. It forms part

CC of haemagglutinin haggA (see also AAAR96030). Harep2 and other haggA

CC repeat unit products (see also AAAR96021 and AAAR96023-24) can be obd.

CC from transformed host cells and used as vaccines to protect humans

CC or animals against periodontal disease. Expression in Salmonella

CC cells allows production of live vaccine. Harep1-4 can also be used

CC to detect the presence of anti-P. gingivalis antibodies and to

CC raise monoclonal antibodies for diagnostic applan.

SO Sequence 456 AA:

Query Match 100.0%: Score 78: DB 17: Length 456:
Best Local Similarity 100.0%: Pred. No. 8-3e-06:
Matches 15: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

QY 1 ALNPNVYLISKDVVG 15

DB 251 ALNPNVYLISKDVVG 265

RESULT 9

AAAR96023
ID AAAR96023 standard; Protein: 456 AA.

AC AAAR96023:

DT 04-SEP-1996 (first entry)

DE P. gingivalis haemagglutinin haggA Harep3 product.

KW Haemagglutinin: haggA; periodontal disease; vaccine; antibody;

KW Harep3.

OS Porphyromonas gingivalis strain 381.

PN M09617936-A2.
 PD 13-JUN-1996.
 XX
 PF 11-DEC-1995; 95WO-US16108.
 XX
 PR 09-DEC-1994; 94US-0353485.
 XX
 PA (UABR-) UAB RES FOUND.
 PA (UYFL) UNIV FLORIDA.
 XX
 PI Han N, Lantz M, Lepine G, Patti JM, Progulskie-Fox A;
 PI Tumwasorn S;
 DR WPI: 1996-287181/29.
 DR N-PSDB; AAT30647.
 XX
 PT Porphyromonas gingivalis genes and proteins - used in the detection
 PT and vaccination against periodontal disease
 PS
 PS Claim 4; Page 110-112; 153pp; English.
 XX
 CC HAREP3 (AAR96023) is the product of the HAREP3 repeat unit (AAT30647)
 CC of the haggA gene (AAT30654) of P. gingivalis 318. It forms part
 CC of haemagglutinin haggA (see also AAR96030). HAREP3 and other haggA
 CC repeat unit products (see also AAR96021-22 and AAR96024) can be obtd.
 CC from transformed host cells and used as vaccines to protect humans
 CC or animals against periodontal disease. Expression in Salmonella
 CC cells allows produ. of live vaccine. HAREP1-4 can also be used
 CC to detect the presence of anti-P. gingivalis antibodies and to
 CC raise monoclonal antibodies for diagnostic appln.
 XX
 SO Sequence 456 AA:
 Query Match 100.0%; Score 78; DB 17; Length 456;
 Best Local Similarity 100.0%; Pred. No. 8.3e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ALNPDNYLISKDVTG 15
 ||||||||||||
 DB 251 ALNPDNYLISKDVTG 265
 RESULT 10
 AAM69490 standard; Protein: 456 AA.
 XX
 AC AAM69490;
 XX
 DT 22-DEC-1998 (first entry)
 XX
 DE Haemagglutinin protein haggA, HAREP3.
 XX
 KM Haemagglutinin protein; periodontal disease; vaccine; haggA.
 KM
 OS Porphyromonas gingivalis.
 OS
 PM US5824791-A.
 PD 20-OCT-1998.
 PD
 PF 11-DEC-1995; 95US-0570311.
 PF
 PR 11-DEC-1995; 95US-0570311.
 PR 08-SEP-1988; 88US-0241640.
 PR 25-JAN-1991; 91US-0647119.
 PR 09-DEC-1994; 94US-0353485.
 XX
 PA (UABR-) UAB RES FOUND.
 PA (UYFL) UNIV FLORIDA.
 XX
 PI Han N, Lantz M, Lepine G, Patti JM, Progulskie-Fox A;
 PI Tumwasorn S;

XX
 DR WPI: 1998-582627/49.
 DR N-PSDB; AAV58877.
 XX
 PT Isolated Porphyromonas gingivalis genes - encoding haemagglutinin
 PT and/or protease poly:peptide(s))
 XX
 PS Claim 1; Column 127-132; 101pp; English.
 XX
 CC This sequence is encoded by a porphyromonas gingivalis gene of the
 CC invention. This sequence represents the haggA haemagglutinin protein. The
 CC polypeptides are used to produce antibodies to organisms associated with
 CC periodontal disease. The antibodies are also used in purification and
 CC identification procedures. The genes and polypeptides are used as
 CC vaccines against periodontal disease.
 XX
 SO Sequence 456-AA:
 Query Match 100.0%; Score 78; DB 19; Length 456;
 Best Local Similarity 100.0%; Pred. No. 8.3e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ALNPDNYLISKDVTG 15
 ||||||||||||
 DB 251 ALNPDNYLISKDVTG 265
 RESULT 11
 AAM69491 standard; Protein: 456 AA.
 XX
 AC AAM69491;
 XX
 DT 22-DEC-1998 (first entry)
 XX
 DE Haemagglutinin protein haggA, HAREP3.
 XX
 KM Haemagglutinin protein; periodontal disease; vaccine; haggA.
 KM
 OS Porphyromonas gingivalis.
 OS
 PM US5824791-A.
 PD 20-OCT-1998.
 PD
 PF 11-DEC-1995; 95US-0570311.
 PF
 PR 11-DEC-1995; 95US-0570311.
 PR 08-SEP-1988; 88US-0241640.
 PR 25-JAN-1991; 91US-0647119.
 PR 09-DEC-1994; 94US-0353485.
 XX
 PA (UABR-) UAB RES FOUND.
 PA (UYFL) UNIV FLORIDA.
 XX
 PI Han N, Lantz M, Lepine G, Patti JM, Progulskie-Fox A;
 PI Tumwasorn S;
 DR WPI: 1998-582627/49.
 DR N-PSDB; AAV58877.
 XX
 PT Isolated Porphyromonas gingivalis genes - encoding haemagglutinin
 PT and/or protease poly:peptide(s))
 XX
 PS Claim 1; Column 133-138; 101pp; English.
 XX
 CC This sequence is encoded by a Porphyromonas gingivalis gene of the
 CC invention. This sequence represents the haggA haemagglutinin protein. The
 CC polypeptides are used to produce antibodies to organisms associated with
 CC periodontal disease. The antibodies are also used in purification and
 CC identification procedures. The genes and polypeptides are used as
 CC vaccines against periodontal disease.

SQ Sequence 456 AA:
 Query Match 100.0%; Score 78; DB 19; Length 456;
 Best Local Similarity 100.0%; Pred. No. 8.3e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ALNPDNYLISKDVTG 15
 ||||||||||||
 DB 251 ALNPDNYLISKDVTG 265
 RESULT 12
 AAR96028
 ID AAR96028 standard; Protein: 1087 AA.
 AC AAR96028:
 DT 04-SEP-1996 (first entry)
 DE P. gingivalis haemagglutinin haGD.
 KM Haemagglutinin; haGD; periodontal disease; vaccine; antibody.
 KN Porphyromonas gingivalis strain FDC381.
 OS MO9617936-A2.
 PN 13-JUN-1996.
 PD 11-DEC-1995; 95WO-US16108.
 PF 09-DEC-1994; 94US-0353485.
 PR (UABR-) UAB RES FOUND.
 PA (UFL) UNIV FLORIDA.
 PI Han N, Lantz M, Lepine G, Patil JM, Proguiske-Fox A;
 PI Tumwasorn S;
 DR MPI: 1996-287181/29.
 DR N-PSDB; AAT30652.
 XX Porphyromonas gingivalis genes and proteins - used in the detection
 PT and vaccination against periodontal disease
 PS Claim 5; Page 65-68; 153pp; English.
 CC P. gingivalis 381 haemagglutinin haGD (AAR96028) was identified as
 CC the product of a gene (AAT30652) isolated from a P. gingivalis 318
 CC genomic library. The haemagglutinin (see also AAR96032) can be obt.
 CC from transformed host cells and used as a vaccine to protect humans
 CC or animals against periodontal disease. Expression in Salmonella
 CC cells allows produ. of a live vaccine. The haemagglutinin can also
 CC be used to detect the presence of anti-P. gingivalis antibodies and
 CC to raise monoclonal antibodies for diagnostic appln.
 XX
 SQ Sequence 1087 AA:
 Query Match 100.0%; Score 78; DB 17; Length 1087;
 Best Local Similarity 100.0%; Pred. No. 2.4e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ALNPDNYLISKDVTG 15
 ||||||||||||
 DB 579 ALNPDNYLISKDVTG 593
 RESULT 13
 AAM69486
 ID AAM69486 standard; Protein: 1087 AA.
 AC AAM69486:
 XX
 XX

DT 22-DEC-1998 (first entry)
 DE Haemagglutinin protein haGD.
 KM Haemagglutinin protein; periodontal disease; vaccine; haGD.
 OS Porphyromonas gingivalis.
 PN US5824791-A.
 PD 20-OCT-1998.
 PF 11-DEC-1995; 95US-0570311.
 PR 11-DEC-1995; 95US-0570311.
 PR 08-SEP-1988; 88US-0241640.
 PR 25-JAN-1991; 91US-0647119.
 PR 09-DEC-1994; 94US-0353485.
 PA (UABR-) UAB RES FOUND.
 PI (UFL) UNIV FLORIDA.
 PI Han N, Lantz M, Lepine G, Patil JM, Proguiske-Fox A;
 PI Tumwasorn S;
 DR MPI: 1998-582627/49.
 DR N-PSDB; AAV58873.
 PT Isolated Porphyromonas gingivalis genes - encoding haemagglutinin
 and/or protease poly(peptide(s))
 PS Claim 1; Column 57-64; 101pp; English.
 CC This sequence is encoded by a Porphyromonas gingivalis gene of the
 CC invention. This sequence represents the haGD haemagglutinin protein. The
 CC polypeptides are used to produce antibodies to organisms associated with
 CC periodontal disease. The antibodies are also used in purification and
 CC identification procedures. The genes and polypeptides are used as
 CC vaccines against periodontal disease.
 XX
 SQ Sequence 1087 AA:
 Query Match 100.0%; Score 78; DB 19; Length 1087;
 Best Local Similarity 100.0%; Pred. No. 2.4e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ALNPDNYLISKDVTG 15
 ||||||||||||
 DB 579 ALNPDNYLISKDVTG 593
 RESULT 14
 AAR96032
 ID AAR96032 standard; Protein: 1358 AA.
 AC AAR96032:
 DT 04-SEP-1996 (first entry)
 DE P. gingivalis haGD haemagglutinin.
 KM Haemagglutinin; haGD; periodontal disease; vaccine; antibody.
 OS Porphyromonas gingivalis strain FDC381.
 PN MO9617936-A2.
 PD 13-JUN-1996.
 PF 11-DEC-1995; 95WO-US16108.
 PR 09-DEC-1994; 94US-0353485.
 XX

PA (UABR-) UAB RES FOUND.
 PA (UYFL) UNIV FLORIDA.
 PI Han N, Lantz M, Lepine G, Patil JM, Proguliske-Fox A;
 PI Tumwasorn S;
 DR WPI: 1996-287181/29.
 DR N-PSDB: AAT30655.
 XX
 PT Porphyromonas gingivalis genes and proteins - used in the detection
 and vaccination against periodontal disease
 XX
 PS Claim 5: Page 125-129; 153pp; English.
 CC
 CC P. gingivalis 381 haemagglutinin hagd (AAR96032) was identified as
 the product of the second open reading frame of the hagd gene
 (AAT30655) derived from P. gingivalis 318 genomic DNA. A first
 CC open reading frame coded for hagd protease (see also AAR96031).
 CC The protease and haemagglutinin can be obtd. from transformed host
 CC cells and used in vaccines to protect humans or animals against
 CC periodontal disease. Expression in Salmonella cells allows produ.
 CC of live vaccines. The haemagglutinin and protease can also be used
 CC to detect the presence of anti-P. gingivalis antibodies and to raise
 CC monoclonal antibodies for diagnostic appln.
 CC
 SQ Sequence 1358 AA;
 XX
 Query Match 100.0%; Score 78; DB 17; Length 1358;
 Best Local Similarity 100.0%; Pred. No. 3.1e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ALNPDNYLISKDVTG 15
 |||||||||||||||
 DB 850 ALNPDNYLISKDVTG 864

RESULT 15
 AAW69494
 ID AAW69494 standard; Protein: 1358 AA.
 AC
 AC AAW69494;
 XX
 DT 22-DEC-1998 (first entry)
 XX
 DE Haemagglutinin protein hagd.
 XX
 KM Haemagglutinin protein; periodontal disease; vaccine; hagd.
 XX
 OS Porphyromonas gingivalis.
 XX
 PN US5824791-A.
 XX
 PD 20-OCT-1998.
 XX
 PE 11-DEC-1995; 95US-0570311.
 XX
 PR 11-DEC-1995; 95US-0570311.
 PR 08-SEP-1988; 88US-0241640.
 PR 25-JAN-1991; 91US-0647119.
 PR 09-DEC-1994; 94US-0353485.
 XX
 PA (UABR-) UAB RES FOUND.
 PA (UYFL) UNIV FLORIDA.
 XX
 PI Han N, Lantz M, Lepine G, Patil JM, Proguliske-Fox A;
 PI Tumwasorn S;
 DR WPI: 1998-582627/49.
 DR N-PSDB: AAV58880.
 XX
 PT Isolated Porphyromonas gingivalis genes - encoding haemagglutinin
 and/or protease poly:peptide(s))
 XX

PS Claim 1: Column 145-158; 101pp; English.
 XX
 CC This sequence is encoded by a Porphyromonas gingivalis gene of the
 CC invention. This sequence represents the hagd haemagglutinin protein. The
 CC polypeptides are used to produce antibodies to organisms associated with
 CC periodontal disease. The antibodies are also used in purification and
 CC identification procedures. The genes and polypeptides are used as
 CC vaccines against periodontal disease.
 CC
 SQ Sequence 1358 AA;
 XX
 Query Match 100.0%; Score 78; DB 19; Length 1358;
 Best Local Similarity 100.0%; Pred. No. 3.1e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ALNPDNYLISKDVTG 15
 |||||||||||||||
 DB 850 ALNPDNYLISKDVTG 864

Search completed: June 26, 2003, 00:51:42
 Job time : 6.29412 secs

Thu Jun 26 11:58:03 2003

us-09-980-370-1.rapb

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 26, 2003, 00:50:39 : Search time 3.17647 Seconds
(without alignments)
510.977 Million cell updates/sec

Title: US-09-980-370-1
Sequence: 1 ALPNDYLVSKDVTG 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 417779 seqs, 108206813 residues
Total number of hits satisfying chosen parameters: 417779

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications-AA:*
1: /cgn2_6/ptodata/1/pubppa/US08_NEW_PUB pep.*
2: /cgn2_6/ptodata/1/pubppa/PT05_NEW_PUB pep.*
3: /cgn2_6/ptodata/1/pubppa/US06_NEW_PUB pep.*
4: /cgn2_6/ptodata/1/pubppa/US06_PUBCOMB pep.*
5: /cgn2_6/ptodata/1/pubppa/US07_NEW_PUB pep.*
6: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB pep.*
7: /cgn2_6/ptodata/1/pubppa/PT05_PUBCOMB pep.*
8: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB pep.*
9: /cgn2_6/ptodata/1/pubppa/US09_NEW_PUB pep.*
10: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB pep.*
11: /cgn2_6/ptodata/1/pubppa/US10_NEW_PUB pep.*
12: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB pep.*
13: /cgn2_6/ptodata/1/pubppa/US60_NEW_PUB pep.*
14: /cgn2_6/ptodata/1/pubppa/US60_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	53.8	2548	10 US-09-851-682A-1	Sequence 1, Appl1
2	41	52.6	166	9 US-09-925-299-1177	Sequence 1177, Ap
3	41	52.6	166	10 US-09-925-299-1177	Sequence 1177, Ap
4	40	51.3	575	9 US-09-738-626-4263	Sequence 4263, Ap
5	40	51.3	1725	9 US-10-037-182-12	Sequence 12, Appl
6	40	51.3	1765	9 US-10-037-182-8	Sequence 8, Appl
7	40	51.3	1786	9 US-10-037-182-6	Sequence 6, Appl
8	40	51.3	1786	9 US-10-037-182-10	Sequence 10, Appl
9	40	51.3	1786	9 US-09-873-676-113	Sequence 113, Appl
10	40	51.3	1786	10 US-09-938-275-6	Sequence 6, Appl1
11	40	51.3	1786	10 US-09-938-275-7	Sequence 7, Appl1
12	39	50.0	72	10 US-09-205-658-178	Sequence 178, App
13	39	50.0	90	10 US-09-205-658-191	Sequence 191, App
14	39	50.0	331	9 US-10-051-643-182	Sequence 182, App
15	39	50.0	331	9 US-09-880-505-182	Sequence 182, App
16	39	50.0	368	10 US-09-801-368-308	Sequence 308, App
17	39	50.0	632	10 US-09-205-658-159	Sequence 159, App
18	39	50.0	636	10 US-09-205-658-160	Sequence 160, App
19	38	48.7	93	10 US-09-731-872-452	Sequence 452, App

20	38	48.7	381	10 US-09-815-242-10962	Sequence 10962, A
21	38	48.7	613	10 US-09-862-027-39	Sequence 39, Appl
22	38	48.7	1694	9 US-10-223-070-19	Sequence 19, Appl
23	37	47.4	355	10 US-09-803-286A-12	Sequence 12, Appl
24	37	47.4	357	10 US-09-947-027-9	Sequence 9, Appl1
25	37	47.4	357	12 US-10-091-009-9	Sequence 9, Appl1
26	37	47.4	822	9 US-09-740-627-11	Sequence 11, Appl
27	37	47.4	822	9 US-10-081-119-16	Sequence 16, Appl
28	36.5	46.8	109	10 US-09-867-550-474	Sequence 474, App
29	36.5	46.8	491	10 US-09-881-752A-128	Sequence 128, App
30	36	46.2	118	9 US-10-106-698-6332	Sequence 6332, Ap
31	36	46.2	120	10 US-09-764-903-62	Sequence 62, Appl
32	36	46.2	135	9 US-09-764-868-830	Sequence 830, App
33	36	46.2	139	9 US-09-901-938-14	Sequence 14, Appl
34	36	46.2	160	10 US-09-764-877-1284	Sequence 1284, Ap
35	36	46.2	168	9 US-10-081-347-22	Sequence 22, Appl
36	36	46.2	168	10 US-09-750-963-13	Sequence 13, Appl
37	36	46.2	181	10 US-09-902-773A-2	Sequence 2, Appl1
38	36	46.2	181	10 US-09-425-021-18	Sequence 18, Appl
39	36	46.2	207	10 US-09-778-927A-65	Sequence 65, Appl
40	36	46.2	243	9 US-10-081-347-21	Sequence 21, Appl
41	36	46.2	243	10 US-09-822-485-15	Sequence 15, Appl
42	36	46.2	243	10 US-09-251-263-2	Sequence 2, Appl1
43	36	46.2	298	10 US-09-801-368-66	Sequence 66, Appl
44	36	46.2	332	9 US-09-984-130-32	Sequence 32, Appl
45	36	46.2	332	9 US-10-215-457-2	Sequence 2, Appl1

ALIGNMENTS

RESULT 1
US-09-851-682A-1
Sequence 1, Application US/09851682A
Patient No. US20020091248A1
GENERAL INFORMATION:
APPLICANT: Adams, Arwen E.
APPLICANT: Chiu, David
APPLICANT: Duhl, David
APPLICANT: Gorman, Susan W.
APPLICANT: Leng, Song
APPLICANT: Shelld, Juliet
TITLE OF INVENTION: MYOSIN IXA AND CYCLIC NUCLEOTIDE GATED
TITLE OF INVENTION: CHANNEL-15 (CNGC-15) POLYNUCLEOTIDES, POLYPEPTIDES,
FILE REFERENCE: 200130.442
CURRENT APPLICATION NUMBER: US/09/851.682A
PRIOR FILING DATE: 2001-05-08
PRIOR APPLICATION NUMBER: US/09/172.422
NUMBER OF SEQ ID NOS: 3
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 1
LENGTH: 2548
TYPE: PRT
ORGANISM: Homo sapien
US-09-851-682A-1
Query Match 53.8%; Score 42; DB 10; Length 2548;
Best Local Similarity 70.0%; Pred. No. 2.3e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
OY 2 LMPNDYLVSK 11
DB 989 LMPNDYLVSK 998
RESULT 2
US-09-925-299-1177
Sequence 1177, Application US/09925299
Publication No. US20030040617A9
GENERAL INFORMATION:

us-09-980-370-1.rapb

Page 2

```

Db          141  ALNOKNFILSEDL 153

RESULT 4
US-09-738-626-4263
: Sequence 4263, Application US/09738626
: Publication No. US20020197605A1
: GENERAL INFORMATION:
: APPLICANT: NAKAGAMA, SATOSHI
: APPLICANT: MIZOGUCHI, HIROSHI
: APPLICANT: ANDO, SEIKO
: APPLICANT: HAYASHI, MIKIRO
: APPLICANT: OCHIAI, KETKO
: APPLICANT: YOKOI, HARUHIKO
: APPLICANT: TATEISHI, NAOKO
: APPLICANT: SENOH, AKIHRO
: APPLICANT: IKEDA, MASATO
: APPLICANT: OZAKI, AKIO
: TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
: FILE REFERENCE: 249-125
: CURRENT APPLICATION NUMBER: US/09/738, 626
: CURRENT FILING DATE: 2000-12-18
: PRIOR APPLICATION NUMBER: JP 99/377484
: PRIOR FILING DATE: 1999-12-16
: PRIOR APPLICATION NUMBER: JP 00/759162
: PRIOR FILING DATE: 2000-04-07
: PRIOR APPLICATION NUMBER: JP 00/780988
: PRIOR FILING DATE: 2000-08-03
: NUMBER OF SEQ ID NOS: 7059
: SOFTWARE: PatentIn ver. 3.0
: SEQ ID NO 4263
:
: LENGTH: 575
:
: TYPE: PRT
: ORGANISM: Corynebacterium glutamicum
US-09-738-626-4263

Query Match      51.3%: Score 40; DB 9; Length 575;
Best Local Similarity 53.3%: Pred. No. 95;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Db          91  ALNRPVRLVPEDETC 105

RESULT 5
US-10-037-182-12
: Sequence 12, Application US/10037182
: Publication No. US20030044899A1
: GENERAL INFORMATION:
: APPLICANT: TRYGGVASON, KARL
: APPLICANT: DOI, MASAYUKI
: APPLICANT: THYBOLL, JILL
: TITLE OF INVENTION: Recombinant Laminin 10
: FILE REFERENCE: 99-274-P
: CURRENT APPLICATION NUMBER: US/10/037,182
: CURRENT FILING DATE: 2001-12-21
: PRIOR APPLICATION NUMBER: 60/257,449
: PRIOR FILING DATE: 2000-12-21
: PRIOR APPLICATION NUMBER: 60/279,282
: PRIOR FILING DATE: 2001-03-28
: NUMBER OF SEQ ID NOS: 36
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 12
:
: LENGTH: 1725
:
: TYPE: PRT
: ORGANISM: Mus musculus
US-10-037-182-12

Query Match      51.3%: Score 40; DB 9; Length 1725;
Best Local Similarity 61.5%: Pred. No. 3,2e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```

QY 2 LNPDNYLISKDVT 14
||||:|
DB 28 LNPDSHLIENVVT 40

RESULT 6

US-10-037-182-6
; Sequence 8, Application US/10037182
; Publication No. US20030044899A1
; GENERAL INFORMATION:
; APPLICANT: Tryggevason, Karl
; APPLICANT: Dol, Masayuki
; APPLICANT: Thyboll, Jili
; TITLE OF INVENTION: Recombinant Laminin 10
; FILE REFERENCE: 99-274-F
; CURRENT APPLICATION NUMBER: US/10/037,182
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/257,449
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/279,282
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 1765
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-037-182-6

Query Match 51.3%; Score 40; DB 9; Length 1765;
Best Local Similarity 61.5%; Pred. No. 3.3e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 LNPDNYLISKDVT 14
||||:|
DB 68 LNPDSHLIENVVT 80

RESULT 7

US-10-037-182-6
; Sequence 6, Application US/10037182
; Publication No. US20030044899A1
; GENERAL INFORMATION:
; APPLICANT: Tryggevason, Karl
; APPLICANT: Dol, Masayuki
; APPLICANT: Thyboll, Jili
; TITLE OF INVENTION: Recombinant Laminin 10
; FILE REFERENCE: 99-274-F
; CURRENT APPLICATION NUMBER: US/10/037,182
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/257,449
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/279,282
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 1786
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-037-182-6

Query Match 51.3%; Score 40; DB 9; Length 1786;
Best Local Similarity 61.5%; Pred. No. 3.3e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 LNPDNYLISKDVT 14
||||:|
DB 89 LNPDSHLIENVVT 101

RESULT 8

US-10-037-182-10

; Sequence 10, Application US/10037182
; Publication No. US20030044899A1
; GENERAL INFORMATION:
; APPLICANT: Tryggevason, Karl
; APPLICANT: Dol, Masayuki
; APPLICANT: Thyboll, Jili
; TITLE OF INVENTION: Recombinant Laminin 10
; FILE REFERENCE: 99-274-F
; CURRENT APPLICATION NUMBER: US/10/037,182
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/257,449
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/279,282
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 1786
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-037-182-10

Query Match 51.3%; Score 40; DB 9; Length 1786;
Best Local Similarity 61.5%; Pred. No. 3.3e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 LNPDNYLISKDVT 14
||||:|
DB 89 LNPDSHLIENVVT 101

RESULT 9

US-09-873-676-113
; Sequence 113, Application US/09873676
; Patient No. US20020077289A1
; GENERAL INFORMATION:
; APPLICANT: Macdonald, Nicholas J.
; APPLICANT: Sim, Kim L.
; TITLE OF INVENTION: Angiostatin and Endostatin Binding Proteins and Methods of Use
; FILE REFERENCE: 05213-0378 (43170-259333)
; CURRENT APPLICATION NUMBER: US/09/873,676
; CURRENT FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: US 60/209,065
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: US 60/289,387
; PRIOR FILING DATE: 2001-05-08
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 113
; LENGTH: 1786
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-873-676-113

Query Match 51.3%; Score 40; DB 10; Length 1786;
Best Local Similarity 61.5%; Pred. No. 3.3e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 LNPDNYLISKDVT 14
||||:|
DB 89 LNPDSHLIENVVT 101

RESULT 10

US-09-938-275-6
; Sequence 6, Application US/09938275
; Patient No. US20020111309A1
; GENERAL INFORMATION:
; APPLICANT: Gerardo Castillo
; APPLICANT: Alan Snow
; TITLE OF INVENTION: Therapeutic and Diagnostic Applications
; TITLE OF INVENTION: of Laminin and Laminin-Derived Protein Fragments

FILE REFERENCE: PROTEO.P03
CURRENT APPLICATION NUMBER: US/09/938.275
CURRENT FILING DATE: 2001-08-16
NUMBER OF SEQ ID NOS: 11
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6
LENGTH: 1786
TYPE: PRF
ORGANISM: Homo Sapiens
PUBLICATION INFORMATION:
DATABASE ACCESSION NUMBER: Swissprot P07942
DATABASE ENTRY DATE: 1988-08-01
US-09-938-275-6

Query Match
Best Local Similarity 51.3%; Score 40; DB 10; Length 1786;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 LNPDMYLSKDYV 14
DB 89 LNPDSHLIENVVT 101

RESULT 11
US-09-938-275-7
Sequence 7, Application US/09938275
Patent No. US20020111309A1
GENERAL INFORMATION:
APPLICANT: Gerardo Castillo
TITLE OF INVENTION: Therapeutic and Diagnostic Applications
TITLE OF INVENTION: of Lamitin and Lamitin-Derived Protein Fragments
FILE REFERENCE: PROTEO.P03
CURRENT APPLICATION NUMBER: US/09/938.275
CURRENT FILING DATE: 2001-08-16
NUMBER OF SEQ ID NOS: 11
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 1786
TYPE: PRF
ORGANISM: Mus Musculus
PUBLICATION INFORMATION:
DATABASE ACCESSION NUMBER: Swissprot P02469
DATABASE ENTRY DATE: 1989-07-01
US-09-938-275-7

Query Match
Best Local Similarity 51.3%; Score 40; DB 10; Length 1786;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 LNPDMYLSKDYV 14
DB 89 LNPDSHLIENVVT 101

RESULT 12
US-09-205-658-178
Sequence 178, Application US/09205658
Patent No. US20010029617A1
GENERAL INFORMATION:
APPLICANT: Ruvkun, Gary
TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR
TITLE OF INVENTION: IMPAIRED GLUCOSE TOLERANCE CONDITIONS
FILE REFERENCE: 00786/351004
CURRENT APPLICATION NUMBER: US/09/205.658
CURRENT FILING DATE: 1998-12-03
EARLIER APPLICATION NUMBER: 08/857,076
EARLIER FILING DATE: 1997-05-15
EARLIER APPLICATION NUMBER: 08/888,534
EARLIER FILING DATE: 1997-07-07
EARLIER APPLICATION NUMBER: US98/10080
EARLIER FILING DATE: 1998-05-15

NUMBER OF SEQ ID NOS: 328
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 178
LENGTH: 72
TYPE: PRF
ORGANISM: Caenorhabditis elegans
US-09-205-658-178

Query Match
Best Local Similarity 50.0%; Score 39; DB 10; Length 72;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 LNPDMYLSKDYV 12
DB 50 MKPDNVLIOKD 60

RESULT 13
US-09-205-658-191
Sequence 191, Application US/09205658
Patent No. US20010029617A1
GENERAL INFORMATION:
APPLICANT: Ruvkun, Gary
TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR
TITLE OF INVENTION: IMPAIRED GLUCOSE TOLERANCE CONDITIONS
FILE REFERENCE: 00786/351004
CURRENT APPLICATION NUMBER: US/09/205.658
CURRENT FILING DATE: 1998-12-03
EARLIER APPLICATION NUMBER: 08/857,076
EARLIER FILING DATE: 1997-05-15
EARLIER APPLICATION NUMBER: 08/888,534
EARLIER FILING DATE: 1997-07-07
EARLIER APPLICATION NUMBER: US98/10080
EARLIER FILING DATE: 1998-05-15
NUMBER OF SEQ ID NOS: 328
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 191
LENGTH: 90
TYPE: PRF
ORGANISM: Caenorhabditis elegans
US-09-205-658-191

Query Match
Best Local Similarity 50.0%; Score 39; DB 10; Length 90;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 LNPDMYLSKDYV 12
DB 68 MKPDNVLIOKD 78

RESULT 14
US-10-051-643-182
Sequence 182, Application US/10051643
Publication No. US20020197265A1
GENERAL INFORMATION:
APPLICANT: Watson, James D.
TITLE OF INVENTION: Methods and Compounds for the Treatment
TITLE OF INVENTION: of Immunologically-Mediated Diseases of the Respiratory
TITLE OF INVENTION: System using Mycobacterium Vaccae
FILE REFERENCE: 11000.100862
CURRENT APPLICATION NUMBER: US/10/051.643
CURRENT FILING DATE: 2002-01-18
PRIOR APPLICATION NUMBER: US09/156,181
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: US 08/996,624
PRIOR FILING DATE: 1997-12-23
NUMBER OF SEQ ID NOS: 208
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 182
LENGTH: 331

```

; TYPE: PRT
; ORGANISM: Mycobacterium vaccae
US-10-051-643-182

```

```

Query Match      50.0%; Score 39; DB 9; Length 331;
Best Local Similarity 63.6%; Pred. No. 76;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

```

OY      4 PDNYLISKDVT 14
      ||:|:|:|  ||
Db      244 PDHYLVSLSVT 254

```

```

RESULT 15
US-09-880-505-182
; Sequence 182, Application US/09880505
; Publication No. US20030007976A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Tan, Paul L.J.
; APPLICANT: Prestidge, Ross
; TITLE OF INVENTION: Methods and Compounds for the Treatment
; FILE REFERENCE: 11000.1007c2
; CURRENT APPLICATION NUMBER: US/09/880,505
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 09/324,542
; PRIOR FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: US 08/997,080
; PRIOR FILING DATE: 1997-12-23
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 182
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Mycobacterium vaccae
US-09-880-505-182

```

```

Query Match      50.0%; Score 39; DB 9; Length 331;
Best Local Similarity 63.6%; Pred. No. 76;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

```

OY      4 PDNYLISKDVT 14
      ||:|:|:|  ||
Db      244 PDHYLVSLSVT 254

```

Search completed: June 26, 2003, 00:54:44
 Job time : 3.17647 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 26, 2003, 00:46:24 : Search time 2.47059 Seconds

(Without alignments)
583.673 Million cell updates/sec

Title: US-09-980-370-1

Sequence: 1 ALNPDNYLISKDVTC 15

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	78	100.0	1526 2	gingipain R (EC 3.4.22.37)
2	78	100.0	1704 2	gingipain R (EC 3.4.22.37)
3	78	100.0	1732 2	lysine-specific cy
4	78	100.0	2628 2	hemagglutinin A -
5	47	60.3	334 2	hypothetical prote
6	45	57.7	265 2	tubulin gamma chain
7	45	57.7	54209 2	cinamyl alcohol d
8	44	56.4	315 2	hypothetical prote
9	44	56.4	639 2	hypothetical prote
10	44	56.4	752 2	phosphoenolpyruvat
11	43	55.1	290 2	kinase-like protei
12	43	55.1	290 2	protein kinase hom
13	42	53.8	318 2	hypothetical prote
14	42	53.8	1777 2	hypothetical prote
15	42	53.8	1777 2	hypothetical prote
16	42	53.8	2548 2	myosin IXA (import
17	41	52.6	146 2	protein kinase hom
18	41	52.6	314 2	growth arrest-spec
19	41	52.6	383 2	hypothetical prote
20	41	52.6	393 2	probable protein k
21	41	52.6	821 2	phosphoenolpyruvat
22	41	52.6	829 2	periplasmic nitrat
23	41	52.6	834 2	Napa periplasmic n
24	40	51.3	883 2	hypothetical prote
25	40	51.3	1786 1	laminin beta-1 cha
26	40	51.3	1786 1	laminin beta-1 cha
27	39.5	50.6	2626 2	myosin-Rhocap prot
28	39.5	50.6	2626 2	phosphoglycerate k
29	39.5	50.6	189 2	hypothetical prote

30	39	50.0	274 2	S28039	nuclease nuca prec
31	39	50.0	274 2	A82523	sugar-non-specific
32	39	50.0	312 2	T32446	hypothetical prote
33	39	50.0	312 2	A89460	protein H42K12.1 f
34	39	50.0	321 2	T24599	hypothetical prote
35	39	50.0	326 2	B46108	outer capsid prote
36	39	50.0	329 2	T46214	hypothetical prote
37	39	50.0	368 2	A54430	hypoxic function t
38	39	50.0	372 1	A54870	cystathionine beta
39	39	50.0	472 1	A54870	cell adhesion rece
40	39	50.0	636 2	G81983	pilin glycosylatio
41	39	50.0	694 2	E69143	hypothetical prote
42	38.5	49.4	13055 2	T16580	hypothetical prote
43	38.5	49.4	302 2	T11675	lactoyglutathione
44	38.5	49.4	917 2	B81309	isooleucine-tRNA 11
45	38	48.7	181 2	T02591	germin-like protei

ALIGNMENTS

RESULT 1

gingipain R (EC 3.4.22.37) precursor - Porphyromonas gingivalis (fragment)

C:Species: Porphyromonas gingivalis

C:Date: 05-Mar-1995 #sequence_revision 12-May-1995 #text_change 31-Mar-1997

C:Accession: S49763

R:Aduse-Opoku, J.; Mull, J.; Slaney, J.M.; Rangarajan, M.; Curtis, M.A.

submitted to the EMBL Data Library, November 1994

A:Description: Cloning, sequence analysis and expression in Escherichia coli of p1r1

A:Reference number: S49763

A:Accession: S49763

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1526 <ADU>

A:Cross-references: EMBL:X82680

C:Genetics:

A:Gene: p1r1

C:Keywords: cysteine proteinase; hydrolase

Query Match 100.0% Score 78; DB 2; Length 1526;

Best Local Similarity 100.0%; Pred. No. 8.8e-05;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1198 ALNPDNYLISKDVTC 1212

RESULT 2

gingipain R (EC 3.4.22.37) precursor - Porphyromonas gingivalis

N:Alternate names: 50K high molecular mass arginine-specific cysteine proteinase: HCP

C:Species: Porphyromonas gingivalis

C:Date: 10-Feb-1995 #sequence_revision 10-Feb-1995 #text_change 08-Oct-1999

C:Accession: A55426; D53113

R:Payloft, N.; Potempa, J.; Pike, R.N.; Prochazka, V.; Kiefer, M.C.; Travis, J.; Barr

J. Biol. Chem. 270, 10077-10100, 1995

A:Title: Molecular cloning and structural characterization of the Arg-gingipain prote

A:Reference number: A55426; MIMD:95138080; PMID:7836351

A:Accession: A55426

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1704 <PAV>

A:Cross-references: GB:015282; NID:9557067; PIDN:AA69539.1; PID:9557068

J. Biol. Chem. 269, 406-411, 1994

A:Title: Lysine- and arginine-specific proteinases from Porphyromonas gingivalis. Iso

A:Reference number: A53113; MIMD:94103245; PMID:8276827

A:Accession: D53113

A:Status: preliminary

A:Molecule type: protein

A:Residues: 228-249 <PIK>

A: Experimental source: H66
A: Note: sequence extracted from NCBI backbone (NCBI:141694)
C: Keywords: cysteine proteinase; hydrolase

Query Match 100.0%; Score 78; DB 2; Length 1704;
Best Local Similarity 100.0%; Pred. No. 9.9e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALNPDNYLSKDYTG 15
DB 1196 ALNPDNYLSKDYTG 1210

RESULT 3

T28651
Lysine-specific cysteine proteinase porphyrin (EC 3.4.22.-) - Porphyromonas gingivalis
N: Alternate names: lysine-specific cysteine proteinase 1, 60k
C: Species: Porphyromonas gingivalis
C: Date: 22-Oct-1999 #sequence-revision 22-Oct-1999 #text-change 17-Nov-2000

C: Accession: T28651; EMBL:AF017059; NID:92738802; PID:92738803; PIDN:AAC26523.1
R: Parkin, C.; Luo, H.; Knight, A.; Ahlquist, J.; Perlman, M.H.
J. Bacteriol. 178, 2734-2741, 1996
A: Title: Analysis of the prtC gene encoding porphyrin, a cysteine proteinase of Porphyromonas gingivalis
A: Reference number: 220899; MUID:96213011; PMID:8631659

A: Accession: T28651
A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: DNA
A: Residues: 1-1732 <GB>
R: Stakkestad, N.; Chai, S.M.; Reynolds, E.C.
Submitted to the EMBL Data Library, October 1996

A: Reference number: 220896
A: Accession: T28651
A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: DNA
A: Residues: 1-795; 1-797-1389; 'N', 1391-1478; 'Y', 1480-1732 <SLA>
A: Cross-references: EMBL:AF017059; NID:92738802; PID:92738803; PIDN:AAC26523.1

A: Cross-references: EMBL:AF017059; NID:92738802; PID:92738803; PIDN:AAC26523.1
R: Parkin, C.; Luo, H.; Knight, A.; Ahlquist, J.; Perlman, M.H.
J. Bacteriol. 178, 2734-2741, 1996
A: Title: Analysis of the prtC gene encoding porphyrin, a cysteine proteinase of Porphyromonas gingivalis
A: Reference number: 220899; MUID:96213011; PMID:8631659

A: Accession: T28651
A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: DNA
A: Residues: 1-1350; 'N', 1352-1363; 'Y', 1365-1447; 'H', 1449-1732 <LEW>
A: Cross-references: EMBL:AF017059; NID:92738802; PID:92738803; PIDN:AAC26523.1

A: Cross-references: EMBL:AF017059; NID:92738802; PID:92738803; PIDN:AAC26523.1
R: Parkin, C.; Luo, H.; Knight, A.; Ahlquist, J.; Perlman, M.H.
J. Bacteriol. 178, 2734-2741, 1996
A: Title: Analysis of the prtC gene encoding porphyrin, a cysteine proteinase of Porphyromonas gingivalis
A: Reference number: 220899; MUID:96213011; PMID:8631659

A: Accession: T28651
A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: DNA
A: Residues: 1-1350; 'N', 1352-1363; 'Y', 1365-1447; 'H', 1449-1732 <LEW>
A: Cross-references: EMBL:AF017059; NID:92738802; PID:92738803; PIDN:AAC26523.1

A: Cross-references: EMBL:AF017059; NID:92738802; PID:92738803; PIDN:AAC26523.1
R: Parkin, C.; Luo, H.; Knight, A.; Ahlquist, J.; Perlman, M.H.
J. Bacteriol. 178, 2734-2741, 1996
A: Title: Analysis of the prtC gene encoding porphyrin, a cysteine proteinase of Porphyromonas gingivalis
A: Reference number: 220899; MUID:96213011; PMID:8631659

A: Accession: T28651
A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: DNA
A: Residues: 1-1350; 'N', 1352-1363; 'Y', 1365-1447; 'H', 1449-1732 <LEW>
A: Cross-references: EMBL:AF017059; NID:92738802; PID:92738803; PIDN:AAC26523.1

A: Cross-references: EMBL:AF017059; NID:92738802; PID:92738803; PIDN:AAC26523.1
R: Parkin, C.; Luo, H.; Knight, A.; Ahlquist, J.; Perlman, M.H.
J. Bacteriol. 178, 2734-2741, 1996
A: Title: Analysis of the prtC gene encoding porphyrin, a cysteine proteinase of Porphyromonas gingivalis
A: Reference number: 220899; MUID:96213011; PMID:8631659

A: Accession: T28651
A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: DNA
A: Residues: 1-1350; 'N', 1352-1363; 'Y', 1365-1447; 'H', 1449-1732 <LEW>
A: Cross-references: EMBL:AF017059; NID:92738802; PID:92738803; PIDN:AAC26523.1

C: Date: 15-Oct-1999 #sequence-revision 15-Oct-1999 #text-change 03-Aug-2001
C: Accession: T28651
R: Han, N.; Whitlock, J.; Proguiske-Fox, A.
Infect. Immun. 64, 4000-4007, 1996
A: Title: The hemagglutinin gene A (haga) of Porphyromonas gingivalis 361 contains four
A: Reference number: 220494; MUID:97047672; PMID:8928061

A: Accession: T28651
A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: DNA
A: Residues: 1-2628 <HAN>
A: Cross-references: EMBL:U01807; NID:91552410; PID:91469916; PIDN:AMB17128.1

A: Accession: T28651
A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: DNA
A: Residues: 1-2628 <HAN>
A: Cross-references: EMBL:U01807; NID:91552410; PID:91469916; PIDN:AMB17128.1

A: Accession: T28651
A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: DNA
A: Residues: 1-2628 <HAN>
A: Cross-references: EMBL:U01807; NID:91552410; PID:91469916; PIDN:AMB17128.1

A: Accession: T28651
A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: DNA
A: Residues: 1-2628 <HAN>
A: Cross-references: EMBL:U01807; NID:91552410; PID:91469916; PIDN:AMB17128.1

A: Accession: T28651
A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: DNA
A: Residues: 1-2628 <HAN>
A: Cross-references: EMBL:U01807; NID:91552410; PID:91469916; PIDN:AMB17128.1

A: Accession: T28651
A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: DNA
A: Residues: 1-2628 <HAN>
A: Cross-references: EMBL:U01807; NID:91552410; PID:91469916; PIDN:AMB17128.1

A: Accession: T28651
A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: DNA
A: Residues: 1-2628 <HAN>
A: Cross-references: EMBL:U01807; NID:91552410; PID:91469916; PIDN:AMB17128.1

A: Accession: T28651
A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: DNA
A: Residues: 1-2628 <HAN>
A: Cross-references: EMBL:U01807; NID:91552410; PID:91469916; PIDN:AMB17128.1

A: Accession: T28651
A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: DNA
A: Residues: 1-2628 <HAN>
A: Cross-references: EMBL:U01807; NID:91552410; PID:91469916; PIDN:AMB17128.1

A: Accession: T28651
A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: DNA
A: Residues: 1-2628 <HAN>
A: Cross-references: EMBL:U01807; NID:91552410; PID:91469916; PIDN:AMB17128.1

A: Accession: T28651
A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: DNA
A: Residues: 1-2628 <HAN>
A: Cross-references: EMBL:U01807; NID:91552410; PID:91469916; PIDN:AMB17128.1

A: Accession: T28651
A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: DNA
A: Residues: 1-2628 <HAN>
A: Cross-references: EMBL:U01807; NID:91552410; PID:91469916; PIDN:AMB17128.1

A: Accession: T28651
A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: DNA
A: Residues: 1-2628 <HAN>
A: Cross-references: EMBL:U01807; NID:91552410; PID:91469916; PIDN:AMB17128.1

A: Accession: T28651
A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: DNA
A: Residues: 1-2628 <HAN>
A: Cross-references: EMBL:U01807; NID:91552410; PID:91469916; PIDN:AMB17128.1

A: Accession: T28651
A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: DNA
A: Residues: 1-2628 <HAN>
A: Cross-references: EMBL:U01807; NID:91552410; PID:91469916; PIDN:AMB17128.1

A: Accession: T28651
A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: DNA
A: Residues: 1-2628 <HAN>
A: Cross-references: EMBL:U01807; NID:91552410; PID:91469916; PIDN:AMB17128.1

A: Accession: T28651
A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: DNA
A: Residues: 1-2628 <HAN>
A: Cross-references: EMBL:U01807; NID:91552410; PID:91469916; PIDN:AMB17128.1

A: Accession: T28651
A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: DNA
A: Residues: 1-2628 <HAN>
A: Cross-references: EMBL:U01807; NID:91552410; PID:91469916; PIDN:AMB17128.1

A: Accession: T28651
A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: DNA
A: Residues: 1-2628 <HAN>
A: Cross-references: EMBL:U01807; NID:91552410; PID:91469916; PIDN:AMB17128.1

A: Accession: T28651
A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: DNA
A: Residues: 1-2628 <HAN>
A: Cross-references: EMBL:U01807; NID:91552410; PID:91469916; PIDN:AMB17128.1

hypothetical protein yscA [imported] - Lactococcus lactis subsp. lactis (strain IL1403)

A;Gene: AT4g19110

Thu Jun 26 11:58:05 2003

us-09-980-370-1.isp

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model
Run on: June 26, 2003, 00:44:19 ; Search time 1.58824 Seconds
(without alignments)
391.721 Million cell updates/sec

Title: US-09-980-370-1
Perfect score: 78
Sequence: 1 ALNDPNYLSKDVTC 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt.40.4

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	78	100.0	989	1 PRTH_PORGI	P46071 porphyromon
2	78	100.0	2628	1 HAGN_PORGI	P46071 porphyromon
3	45	57.7	265	1 TBG_COCH	P46071 porphyromon
4	45	57.7	313	1 TBG_COCH	P46071 porphyromon
5	45	57.7	313	1 TBG_COCH	P46071 porphyromon
6	41	52.6	314	1 TBG2_MOUSE	P46071 porphyromon
7	41	52.6	462	1 TBG2_MOUSE	P46071 porphyromon
8	41	52.6	618	1 CHAN_GLOBI	P46071 porphyromon
9	40	51.3	1786	1 LMB1_HUMAN	P46071 porphyromon
10	40	51.3	1786	1 LMB1_HUMAN	P46071 porphyromon
11	39	50.0	274	1 NUC1_MOUSE	P46071 porphyromon
12	39	50.0	274	1 NUC1_MOUSE	P46071 porphyromon
13	39	50.0	368	1 ROX1_YEAST	P46071 porphyromon
14	39	50.0	471	1 CD36_HUMAN	P46071 porphyromon
15	38.5	49.4	302	1 LGUL_SCHPO	P46071 porphyromon
16	38.5	49.4	851	1 MCEL_FOVY	P46071 porphyromon
17	38.5	49.4	917	1 S11_GAMME	P46071 porphyromon
18	38.5	49.4	917	1 S11_GAMME	P46071 porphyromon
19	38.5	49.4	917	1 S11_GAMME	P46071 porphyromon
20	38	48.7	369	1 CBR1_HUMAN	P46071 porphyromon
21	38	48.7	418	1 METR_HAEN	P46071 porphyromon
22	38	48.7	418	1 METR_HAEN	P46071 porphyromon
23	38	48.7	459	1 MB23_ARATH	P46071 porphyromon
24	38	48.7	484	1 TRPC_YEAST	P46071 porphyromon
25	38	48.7	613	1 TRPC_YEAST	P46071 porphyromon
26	38	48.7	726	1 AMSA_BMAN	P46071 porphyromon
27	38	48.7	781	1 YG5A_YEAST	P46071 porphyromon
28	37.5	48.1	705	1 CWBA_BACU	P46071 porphyromon
29	37	47.4	147	1 LEC2_ANTOR	P46071 porphyromon
30	37	47.4	321	1 NMDA_SULTO	P46071 porphyromon
31	37	47.4	321	1 NMDA_SULTO	P46071 porphyromon
32	37	47.4	321	1 NMDA_SULTO	P46071 porphyromon
33	37	47.4	452	1 PR11_PLATK	P46071 porphyromon

ALIGNMENTS

RESULT 1
ID PRTH_PORGI STANDARD: PRT: 989 AA.

34	37	47.4	492	1 VAR2_ACEAT	Q36690 acetabulari
35	37	47.4	504	1 CK13_YEAST	P39962 saccharomyc
36	37	47.4	677	1 HSF_KLJUA	P22121 kluyveromyc
37	37	47.4	681	1 YDHE_SCHPO	Q92359 schizosacch
38	37	47.4	761	1 AFG3_YEAST	P39925 saccharomyc
39	37	47.4	818	1 PPSA_SYNY3	Q55905 synchocyst
40	37	47.4	818	1 PPSA_SYNY3	P41116 saccharomyc
41	37	47.4	818	1 PPSA_SYNY3	Q52350 saccharomyc
42	37	47.4	818	1 PPSA_SYNY3	P23654 drosophila
43	37	47.4	846	1 NRT_DROME	P20289 rattus norv
44	37	47.4	846	1 NRT_DROME	P20289 rattus norv
45	36	46.2	177	1 VEG1_RAT	P20289 rattus norv

AC P46071: 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE Porphyromonas gingivalis (Bacteroides; Bacteroidales;
GN Porphyromonadaceae; Porphyromonadaceae; Porphyromonadaceae;
OC Bacteroides; Bacteroidales; Bacteroidales; Bacteroidales;
OC Porphyromonadaceae; Porphyromonadaceae; Porphyromonadaceae;
OC NCP1_taxid=8377;
RP STRAIN=WB3;
RC MEDLINE=95012612; PubMed=7927685;
RA Fletcher H.M., Schenkeln H.A., Macrina F.L.;
RT Cloning and characterization of a new protease gene (prth) from
RT Porphyromonas gingivalis (Bacteroides; Bacteroidales; Bacteroidales;
RP Immun. 62:479-486(1994).
RP ERRATUM.
RA Fletcher H.M., Schenkeln H.A., Macrina F.L.;
RL Infect. Immun. 62:5707-5707(1994).
CC -1- FUNCTION: CLEAVES HUMAN COMPLEMENT COMPONENT C3. MAY ENHANCE
CC IMMUNE RESPONSE. PLAYS AN IMPORTANT ROLE IN SOFT TISSUE INFECTIONS
CC AND IS A VIRULENCE FACTOR.
CC -1- SUBCELLULAR LOCATION: IN MEMBRANE VESICLES.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C25.
CC THIS SWISS-PROT entry is copyright. It is produced through collaboration
CC between the EMBL, EMBL, EMBL, EMBL, EMBL, EMBL, EMBL, EMBL, EMBL, EMBL,
CC and the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC EMBL: J27483; J27483; J27483; J27483; J27483; J27483; J27483; J27483; J27483; J27483;
CC HSP: P23882; JFWT.
DR MEROPS: C25.001; -;
DR InterPro: IPR001769; Peptidase_C25.
DR InterPro: IPR002376; Formyl_transf.
DR Pfam: PF00551; Formyl_transf. 1.
DR Pfam: PF01361; Peptidase_C25. 1.
FT REPEAT 270 323
FT REPEAT 528 581
SO SEQUENCE 989 AA: 110238 MW: 7855FEBA3AC8944C CRC64:
Query Match 100.0% Score 78: DB 1: Length 989:
Matches 15: Conservative 0: Mismatches 0: Gaps 0:

```

OY      1 ALPNPNYLISKDVTG 15
DB      149 ALPNPNYLISKDVTG 163

RESULT 2
HAGA_PORCI STANDARD: PRT: 2628 AA.
AC 051845..2000 (rel. 39, Created)
AD 051845..2000 (rel. 39, Last sequence update)
DT 30-MAY-2000 (rel. 39, Last annotation update)
DE Hemagglutinin A precursor.
GN HAGA.
OS Porphyromonas gingivalis (Bacteroides gingivalis),
OC Bacteria; Bacteroidetes; Bacteroidales; Bacteroidales;
OC Bacteroidales; Bacteroidales; Bacteroidales;
OX NCBI_TaxID=837;

RP SEQUENCE FROM N.A.
RX STRAIN=381;
RC MEDLINE=97047672; PubMed=8926061;
RT "The hemagglutinin (HAGA) of Porphyromonas gingivalis 381
RL Infect. Immun. 64:4000-4007(1996).
CC CC
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C25.
CC THIS SWISS-PROT entry is produced through collaboration
CC between the Swiss Institute of Bioinformatics and the
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U01807; A017128.1;
DR InterPro: IPR001769; Peptidase_C25.
KM Hemagglutinin; Virulence; Hydrolyase; Signal; Repeat.
FT SIGNAL 25 2628
FT DOMAIN 25 539
FT DOMAIN 540 995
FT DOMAIN 996 1451
FT DOMAIN 1452 1907
FT DOMAIN 2024 263324 MW: 6104052540C950A CRC64:
SO SEQUENCE 2628 AA; 263324 MW: 6104052540C950A CRC64:

Query Match
Best Local Similarity 100.0%; Score 78; DB 1; Length 2628;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 ALPNPNYLISKDVTG 15
DB      744 ALPNPNYLISKDVTG 758

RESULT 3
TREG_COCHIE STANDARD: PRT: 265 AA.
AD 01-FEB-1995 (rel. 31, Created)
DT 01-NOV-1995 (rel. 32, Last sequence update)
DT 15-JUN-2002 (rel. 41, Last annotation update)
DE Tubulin gamma chain (gamma tubulin) (Fragmant).
OS Cechlobolus neteroslophus (Drechslera maydis).
OC Fungi; Basidiomycota; Basidiomycetes; Basidiomycetes;
OC Pleosporales; Pleosporales; Cestibolales;
OX NCBI_TaxID=5016;

RP SEQUENCE FROM N.A.
RX STRAIN=381;
RC MEDLINE=97047672; PubMed=8926061;
RT "The hemagglutinin (HAGA) of Porphyromonas gingivalis 381
RL Infect. Immun. 64:4000-4007(1996).
CC CC
CC -1- SIMILARITY: BELONGS TO THE TUBULIN FAMILY.
CC THIS SWISS-PROT entry is produced through collaboration
CC between the Swiss Institute of Bioinformatics and the
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X74455; CA52464.1;
DR InterPro: IPR000217; Tubulin.
DR InterPro: IPR003008; Tubulin_FtsZ.
KM PROSITE: P800227; TUBULIN: 1.
RC Microtubules; GTP-binding.
FT NP_BIND 77 83
FT NON_TER 265 265 GTP (POTENTIAL).
SO SEQUENCE 265 AA; 29567 MW: ASDAC23E7D62DC6 CRC64:

Query Match
Best Local Similarity 61.5%; Score 45; DB 1; Length 265;
Matches 0; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY      3 NPNPNYLISKDVTG 15
DB      22 NPNPNYLISKDVTG 34

RESULT 4
CADH_MEDSA STANDARD: PRT: 358 AA.
AC 01-JUL-1993 (rel. 26, Created)
DT 01-JUL-1993 (rel. 26, Last sequence update)
DT 15-JUN-2002 (rel. 41, Last annotation update)
DE Cinnamyl-alcohol dehydrogenase (EC 1.1.1.195) (CAD).
GN CAD2.
OS Medicago sativa (Alfalfa).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Cistaceae; Fabaceae; Papilionoideae; Trifoliales; Medicago.
OX NCBI_TaxID=3879;

RP SEQUENCE FROM N.A.
RX STRAIN=cv. Apollo;
RC van Doorslaere J., Baucher M., Feuillet C., Boudet A.M.,
RC "Cloning of cinnamyl alcohol dehydrogenase cDNAs from two important
RC economic species: alfalfa and poplar. Demonstration of a high homology
RT of the gene within angiosperms."
RL Plant Physiol. Biochem. 33:105-109(1995).
RN
RP SEQUENCE FROM N.A. AND CHARACTERIZATION.
RN MEDLINE=20044095; PubMed=1057854;
RX Brull E.M., Abrahams S., Hayes C.M., Jenkins C.L., Watson J.M.;
RT "Molecular characterization and expression of a wound-inducible cDNA
RT encoding a novel cinnamyl-alcohol dehydrogenase enzyme in lucerne
RT (Medicago sativa L.)."
RN Plant Mol. Biol. 11:29-39(1990)
RN "Purification and characterization of the final step in a branch of
RN the prenylpropanoid synthesis specific for production of lignin
RN monomers. It acts on coniferyl-, sinapyl-, 4-coumaryl- and
RN cinnamyl-alcohol."

```

CC -1- CATALYTIC ACTIVITY: Cinnamyl alcohol + NADP(+) - cinnamaldehyde +
CC NADPH.
CC -1- COFACTOR: ZINC (BY SIMILARITY).
CC -1- PATHWAY: Lipid synthesis.
CC -1- TISSUE SPECIFICITY: MOST ACTIVELY EXPRESSED IN STEM, HYPOTHALAMUS AND
CC ROOT TISSUE.
CC -1- SIMILARITY: BELONGS TO THE ZINC-CONTAINING ALCOHOL DEHYDROGENASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Z19573; CAA79625.1; -
CC EMBL: AF083332; AAC35845.1; -
CC PIR: S31572; S31572.
CC InterPro: IPR002328; ADH_zinc.
CC InterPro: IPR002085; Adh_zn_family.
CC Pfam: PF00107; adh_zinc.1.
CC PROSITE: PS00059; ADH_ZINC.1.
CC Oxidoreductase; NADP; Zinc; Lipid biosynthesis.
CC METAL 48 48 ZINC (CATALYTIC) (BY SIMILARITY).
CC METAL 70 70 ZINC (CATALYTIC) (BY SIMILARITY).
CC METAL 101 101 ZINC (SECOND ATOM) (BY SIMILARITY).
CC METAL 104 104 ZINC (SECOND ATOM) (BY SIMILARITY).
CC METAL 107 107 ZINC (SECOND ATOM) (BY SIMILARITY).
CC METAL 115 115 ZINC (SECOND ATOM) (BY SIMILARITY).
CC METAL 164 164 ZINC (CATALYTIC) (BY SIMILARITY).
CC SO SEQUENCE 358 AA; 38948 MW; FBA609408D01B56 CRC64;
Oy 2 LNPNDYLSKDYTG 15
Db 225 LGADNYLVSDTVG 238
Query Match 57.7%; Score 45; DB 1; Length 358;
Best Local Similarity 57.1%; Pred. No. 2.9;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
RESULT 5
GAS2_HUMAN STANDARD; PRT; 313 AA.
ID GAS2_HUMAN
AC 043903;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Growth-arrest-specific protein 2 (GAS-2).
GN GAS2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A.
RP TISSUE=Kidney;
RX MEDLINE=98190528; PubMed=9521882;
RA Collavito L., Buzza M., Saccone S., Bernard L., Federico C.,
RA della Valle G., Brancolini C., Schneider C.;
RT "cDNA characterization and chromosome mapping of the human GAS2
RT gene".
RL Genomics 48:265-269(1998).
-1- FUNCTION: MAY PLAY A ROLE IN APOPTOSIS BY ACTING AS A CELL DEATH
CC SUBSTRATE FOR CASPASES. IS CLEAVED DURING APOPTOSIS AND THE
CC CLEAVED FORM INDUCES DRAMATIC REARRANGEMENTS OF THE ACTIN
CC CYTOSKELETON AND POTENT CHANGES IN THE SHAPE OF THE AFFECTED
CC CELLS. MAY BE INVOLVED IN THE MEMBRANE RUFLING PROCESS (BY
CC SIMILARITY).
-1- SUBCELLULAR LOCATION: COMPONENT OF THE MICROFILAMENT SYSTEM.
CC COLOCALIZES WITH ACTIN FIBERS AT THE CELL BORDER AND ALONG THE

CC STRESS FIBERS IN GROWTH-ARRESTED FIBROBLASTS. MAINLY
CC MEMBRANE-ASSOCIATED. WHEN HYPERPHOSPHORYLATED, ACCUMULATES AT
CC MEMBRANE RUFLINGS (BY SIMILARITY).
CC -1- TISSUE SPECIFICITY: UNIDIRECTIONALLY EXPRESSED WITH HIGHEST LEVELS IN
CC LIVER, LUNG, AND KIDNEY. NOT FOUND IN SPLEEN.
CC -1- DEVELOPMENTAL STAGE: SPECIFICALLY EXPRESSED AT GROWTH ARREST.
CC -1- PTM: CLEAVED, DURING APOPTOSIS, ON A SPECIFIC ASPARTIC RESIDUE BY
CC CASPASES.
CC -1- PTM: PHOSPHORYLATED ON SERINE RESIDUES DURING THE G0-G1 TRANSITION
CC PHASE (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE GAS2 FAMILY.
CC -1- SIMILARITY: CONTAINS 1 CALPONIN-HOMOLOGY (CH) DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U95032; AAC52058.1; -
CC EMBL: HGNC:4167; GAS2.
CC PIR: M62835; -
CC InterPro: IPR001715; Calponin-like.
CC InterPro: IPR003108; GAS2.
CC Pfam: PF02187; GAS2.1.
CC SMART: SM00033; CH.1.
CC SMART: SM00243; GAS2.1.
CC PROSITE: PS00021; CH.1.
CC KW Growth arrest; Phosphorylation; Apoptosis; Cell cycle;
CC Glycoskeleton.
CC FT DOMAIN 34 156 CH.
CC FT SITE 278 279 APOPTOSIS.
CC SO SEQUENCE 313 AA; 34945 MW; A5FBA611A0C8B36F CRC64;
Oy 2 LNPNDYLI 9
Db 294 MNPNDYLV 301
Query Match 52.6%; Score 41; DB 1; Length 313;
Best Local Similarity 75.0%; Pred. No. 12;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
RESULT 6
GAS2_MOUSE STANDARD; PRT; 314 AA.
ID GAS2_MOUSE
AC P11862;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Growth-arrest-specific protein 2 (GAS-2).
GN GAS2 OR GAS-2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=88311081; PubMed=3409319;
RA Schneider C., King R.M., Philipson L.;
RT "Genes specifically expressed at growth arrest of mammalian cells".
RL Cell 54:787-793(1988).
[2]
RN CHARACTERIZATION.
RX MEDLINE=92299682; PubMed=1607387;
RA Brancolini C., Bottega S., Schneider C.;
RT "GAS2, a growth arrest-specific protein, is a component of the
RT microfilament network system".
RL J. Cell Biol. 117:1251-1261(1992).
[3]

RP PHOSPHORYLATION, AND SUBCELLULAR LOCATION.
 RX MEDLINE-9416511; PubMed-812096;
 RA Brancolini C., Schneider C.;
 RT "phosphorylation of the growth arrest-specific protein Gas2 is
 RT coupled to actin rearrangements during G0->G1 transition in
 RT NIH 3T3 cells";
 RL J. Cell Biol. 124:743-756(1994).
 RN [4]
 RP PROCESSING
 RX MEDLINE-96080153; PubMed-7489707;
 RA Brancolini C., Benedetti M., Schneider C.;
 RT "Microfilament reorganization during apoptosis: The role of Gas2, a
 RT possible substrate for ICE-like proteases";
 RL EMBJ J. 14:5179-5190(1995).
 RN [5]
 RP FUNCTION AND DEVELOPMENTAL STAGE.
 RX MEDLINE-99160524; PubMed-10049561;
 RA Lee K.K., Tang M.K., Yew D.T., Chow P.H., Yee S.P., Schneider C.,
 RT Brancolini C.;
 RT "gas2 is a multifunctional gene involved in the regulation of
 RT apoptosis and chondrogenesis in the developing mouse limb.";
 RL Dev. Biol. 207:14-25(1999).
 CC -1- FUNCTION: MAY PLAY A ROLE IN APOPTOSIS BY ACTING AS A CELL DEATH
 CC SUBSTRATE FOR CASPASES. IS CLEAVED DURING APOPTOSIS AND THE
 CC CLEAVED FORM INDUCES DRAMATIC REARRANGEMENTS OF THE ACTIN
 CC CYTOSKELETON AND POTENT CHANGES IN THE SHAPE OF THE AFFECTED
 CC CELLS. MAY PLAY A ROLE IN CHONDROCYTE PROLIFERATION AND
 CC DIFFERENTIATION, AND IN LIMB MYOGENESIS. MAY BE INVOLVED IN THE
 CC REGULATION OF THE APOPTOSIS IN THE INTERDIGITAL TISSUES OF THE
 CC DEVELOPING HINDLIMB. MAY BE INVOLVED IN THE MEMBRANE RUFLING
 CC PROCESS.
 CC -1- SUBCELLULAR LOCATION: COMPONENT OF THE MICROFILAMENT SYSTEM.
 CC COLOCALIZES WITH ACTIN FIBERS AT THE CELL BORDER AND ALONG THE
 CC STRESS FIBERS IN GROWTH-ARRESTED FIBROBLASTS. MAINLY
 CC MEMBRANE-ASSOCIATED. WHEN HYPERPHOSPHORYLATED, ACCUMULATES AT
 CC MEMBRANE RUFLINGS.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN MOST TISSUES. HIGHEST LEVELS IN
 CC LIVER, LUNG AND KIDNEY. IN THE EMBRYO STRONGLY EXPRESSED IN
 CC REGIONS THAT UNDERGO EXTENSIVE APOPTOSIS, SUCH AS THE
 CC INTERDIGITAL TISSUES, THE CRANIOFACIAL MESenchyme AND THE
 CC CARTRIDGE OF THE LIMBS.
 CC -1- DEVELOPMENTAL STAGE: AT EMBRYONIC DAYS E11.5 AND E13.5 STRONGLY
 CC EXPRESSED IN THE SOFT CONNECTIVE TISSUE OF THE FACE AND TRUNK, AND
 CC IN THE INVERTERTEBRAL TISSUES. LOW LEVELS ARE FOUND IN BRAIN AND
 CC NEURAL TUBE. LOW LEVELS ARE FOUND IN DAY E13.5 LUNG, KIDNEY, EYE
 CC LENS AND IN VERTEBRAL CARTILAGE LOCATED CRANIALLY. IN DAY E11.5
 CC HINDLIMBS WEAKLY EXPRESSED BY THE MESenchymal CELLS SURROUNDING
 CC THE PERSPECTIVE CARTILAGE-FORMING REGIONS. IN DAY E12.5 HINDLIMBS
 CC STRONGLY EXPRESSED BY CELLS ENVELOPING THE CHONDROGENIC PRIMORDIA
 CC OF THE DIGITS, METATARSALS, TIBIA, AND FEMUR, AND THE SOFT
 CC CONNECTIVE TISSUE IN THE INTERDIGITAL TISSUES. IN DAY E13.5
 CC HINDLIMBS EXPRESSION IS MAINTAINED IN THE INTERDIGITAL TISSUES
 CC LOCATED PROXIMALLY AND IS FOUND IN SOME CHONDROCYTES IN THE
 CC STYLOPOD AND IN MESenchymal CELLS SURROUNDING THE CARTILAGE IN THE
 CC AUTOPOD AND ZYGPOPOD. IN DAY 13.5 FORELIMB STRONGLY EXPRESSED IN
 CC THE PRE-HYPERTROPIC AND HYPERTROPIC REGIONS OF THE HUMERUS.
 CC RADIUS, AND ULNA. EXPRESSION IN HYPERTROPIC CHONDROCYTES IS
 CC MAINTAINED AT DAY E14.5 AND IS NOT DETECTABLE AT DAY E15.5. AT DAY
 CC 14.5 ALSO EXPRESSED BY CHONDROCYTES IN THE CARTILAGE FORMING THE
 CC CARPALS AND TARSAIS AND BY MESenchymal CELLS IN THE PROCESS OF
 CC CONDENSING TO FORM TENDONS. IN DAY E13.5 HINDLIMBS EXPRESSED IN
 CC SOME MYOBLASTS IN THE PROXIMAL MYOGENIC REGION. IN OLDER LIMBS
 CC EXPRESSION IS MAINTAINED IN THE MYOTUBES.
 CC -1- INDUCTION: DOWN-REGULATED BY MITOGENS.
 CC -1- PTM: CLEAVED DURING APOPTOSIS, ON A SPECIFIC ASPARTIC RESIDUE BY
 CC CASPASES (BY SIMILARITY).
 CC -1- PTM: PHOSPHORYLATED ON SERINE RESIDUES DURING THE G0-G1 TRANSITION
 CC PHASE.
 CC -1- SIMILARITY: BELONGS TO THE GAS2 FAMILY.
 CC -1- SIMILARITY: CONTAINS 1 CALPONIN-HOMOLOGY (CH) DOMAIN.
 CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-2 IS THE INITIATOR.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb.sib.ch).
 CC -----
 CC EMBL: M21828; AAA37660.1; -
 CC DR PIR: A31580; A31590
 CC DR MGD: MGI:95657; Gas2
 CC DR InterPro: IPR001715; Calponin-like.
 CC DR InterPro: IPR003108; GAS2.
 CC DR Pfam: PF02187; GAS2; 1.
 CC DR SMART: SM00033; CH; 1.
 CC DR SMART: SM00243; GAS2; 1.
 CC DR PROSITE: PS50021; CH; 1.
 CC KM Growth arrest; Phosphorylation; Apoptosis; Cell cycle;
 CC KX Cytoskeleton.
 CC FT DOMAIN 35 157 CH.
 CC FT SITE 279 280 CLEAVAGE (BY A CASPASE DURING
 CC FT MUTAGEN 279 279 APOPTOSIS).
 CC FT D->A: ABOLISHES PROTEOLYTIC PROCESSING.
 CC SQ SEQUENCE 314 AA; 34900 MW; 7F2CC704B4057FAC CRC64;
 CC
 CC Query Match 52.6%; Score 41; DB 1; Length 314;
 CC Best Local Similarity 75.0%; Pred. No. 12;
 CC Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC QY 2 LNPNYIL 9
 CC DB 295 MNPNYLY 302
 CC
 CC RESULT 7
 CC TBG2_EUPCR STANDARD: PRT: 462 AA.
 CC ID P54404;
 CC DT 01-OCT-1996 (Rel. 34, Created)
 CC DT 01-OCT-1996 (Rel. 34, Last sequence update)
 CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
 CC DE Tubulin gamma-2 chain (Gamma-2 tubulin).
 CC OS Euplates crassus.
 CC OC Eukaryota; Alveolata; Ciliophora; Spirotrichea; Hypotrichia;
 CC CC Euplotida; Euplotidae; Moneuplotes.
 CC OX NCBI_TaxID=5936;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RX MEDLINE-98192540; PubMed-9524221;
 CC RA Tan M., Heckmann K.;
 CC RT "The two gamma-tubulin-encoding genes of the ciliate Euplates crassus
 CC differ in their sequences, codon usage, transcription initiation
 CC sites and poly(A) addition sites";
 CC RL Gene 210:53-60(1998).
 CC CC -1- FUNCTION: TUBULIN IS THE MAJOR CONSTITUENT OF MICROTUBULES.
 CC GAMMA TUBULIN IS FOUND AT MICROTUBULE ORGANIZING CENTERS (MTOC)
 CC SUCH AS THE SPINDLE POLES OR THE CENTROSOME, SUGGESTING THAT IT
 CC IS INVOLVED IN THE MINUS-END NUCLEATION OF MICROTUBULE ASSEMBLY.
 CC -1- SIMILARITY: BELONGS TO THE TUBULIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb.sib.ch).
 CC -----
 CC EMBL: X85235; CA59490.1; -
 CC DR EMBL: Y09551; CA470742.1; -
 CC DR InterPro: IPR0000217; Tubulin.
 CC DR InterPro: IPR003008; Tubulin_FtsZ.
 CC DR Pfam: PF00091; tubulin_1.
 CC DR PRINTS: PRO1161; TUBULIN.

DR PROSITE: PS00227; TUBULIN: 1.
 KW Microtubules; GTP-binding; Multigene family.
 FT NP_BIND 142 148 GTP (POTENTIAL).
 SO SEQUENCE 462 AA: 52117 MW: 56387559AC4846 CRC64;
 Query Match 52.6%; Score 41; DB 1; Length 462;
 Best Local Similarity 46.2%; Pred. No. 19;
 Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 3 NPDNVLISKDVYTG 15
 11:1:11:1:1
 Db 87 NPDNVLISKDVYTG 99

RESULT 8
 CHAA_CLOBI STANDARD: PRT: 618 AA.
 ID CHAA_CLOBI
 AC 005102;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Pesticidal crystal-like protein cry17Aa (insecticidal toxin
 DE CRYXVIIAa) (CDB72 mosquito toxin).
 GN CRY17AA OR CRYXVIIAa) OR CDB72.
 OS Clostridium bifermentans.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 CC Clostridium.
 CC NCBI_Taxid=1490;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CH18 / subsp. Malaysia;
 RX MEDLINE=98267211; PubMed=9602158;
 RA Barloy F., Lecadet M.-M., Delecluse A.;
 RT Cloning and sequencing of three new putative toxin genes from
 RL Clostridium bifermentans CH18.";
 Gene 211:293-299(1998)
 CC -1- FUNCTION: NOT SIGNIFICANTLY TOXIC TO MOSQUITO LARVAE MAY INCREASE
 THE TOXICITY OF THE PESTICIDIAL CRYSTAL-LIKE PROTEIN CRY16A
 (CBMT1).
 CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).
 CC -1- DEVELOPMENTAL STAGE: PRODUCED DURING SPOULATION.
 CC -1- SIMILARITY: BELONGS TO THE DELTA ENDOTOXIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: X99478; CAA67841.1;
 DR InterPro: IPR001178; Endotoxin.
 DR Pfam: PF00555; endotoxin; 1.
 KW Toxin; Sporulation.
 FT DOMAIN 174 179
 FT SEQUENCE 618 AA: 71651 MW: 1B9C5251B9AD329D CRC64;
 Query Match 52.6%; Score 41; DB 1; Length 618;
 Best Local Similarity 50.0%; Pred. No. 26;
 Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 1 ALNPDNVLISKD 12
 11:1:11:1:1
 Db 436 SISPNYLFKDKD 447

RESULT 9
 LMB1_HUMAN STANDARD: PRT: 1786 AA.
 ID LMB1_HUMAN
 AC P07942;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Laminin beta-1 chain precursor (Laminin B1 chain).
 GN LMB1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90368768; PubMed=1975589;
 RA Voelteenaho R., Chow L.T., Tryggvason K.;
 RT Structure of the human laminin B1 chain gene.";
 RL J. Biol. Chem. 265:15611-15616(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87280097; PubMed=3611077;
 RA Pihlajaniemi T., Eddy R., Fukushima Y., Byers M., Shows T.,
 RA Pihlajaniemi T., Saraste M., Tryggvason K.;
 RT "Human laminin B1 chain. A multidomain protein with gene (LMB1)
 RT locus in the q22 region of chromosome 7.";
 RL J. Biol. Chem. 262:10454-10462(1987).
 RN [3]
 RP SEQUENCE OF 1276-1709 FROM N.A.
 RX MEDLINE=88021029; PubMed=3661559;
 RA Jaye M., Modi W.S., Ricca G.A., Mudd R., Chiu I.M., O'Brien S.J.,
 RA Drohan W.N.;
 RT "Isolation of a cDNA clone for the human laminin-B1 chain and its
 RT gene localization.";
 RL Am. J. Hum. Genet. 41:605-615(1987).
 CC -1- FUNCTION: Binding to cells via a high affinity receptor, laminin
 CC is thought to mediate the attachment, migration, and organization
 CC of cells into tissues during embryonic development by interacting
 CC with other extracellular matrix components.
 CC -1- SUBUNIT: Laminin is a complex glycoprotein, consisting of three
 CC different polypeptide chains (alpha, beta, gamma), which are bound
 CC to each other by disulfide bonds into a cross-shaped molecule
 CC comprising one long and three short arms with globules at each
 CC end.
 CC THE BETA-1 CHAIN IS A SUBUNIT OF LAMININ-1 (EHS LAMININ), LAMININ-
 CC 2 (MEROSIN), AND LAMININ-6 (K-LAMININ).
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- TISSUE SPECIFICITY: FOUND IN THE BASEMENT MEMBRANES (MAJOR
 CC COMPONENT).
 CC -1- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT
 CC WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.
 CC -1- SIMILARITY: CONTAINS 1 LAMININ N-TERMINAL DOMAIN (DOMAIN VI).
 CC -1- SIMILARITY: CONTAINS 12.5 LAMININ EGF-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 LAMININ DOMAIN IV.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: M61951; AAA59486.1;
 DR EMBL: M58147; AAA59486.1; JOINED.
 DR EMBL: M61917; AAA59486.1; JOINED.
 DR EMBL: M61918; AAA59486.1; JOINED.
 DR EMBL: M61921; AAA59486.1; JOINED.
 DR EMBL: M61922; AAA59486.1; JOINED.
 DR EMBL: M61923; AAA59486.1; JOINED.
 DR EMBL: M61924; AAA59486.1; JOINED.
 DR EMBL: M61925; AAA59486.1; JOINED.
 DR EMBL: M61926; AAA59486.1; JOINED.
 DR EMBL: M61927; AAA59486.1; JOINED.
 DR EMBL: M61928; AAA59486.1; JOINED.
 DR EMBL: M61929; AAA59486.1; JOINED.
 DR EMBL: M61930; AAA59486.1; JOINED.
 DR EMBL: M61931; AAA59486.1; JOINED.

DR EMBL: M61932: AAA59486.1: JOINED.
 DR EMBL: M61933: AAA59486.1: JOINED.
 DR EMBL: M61934: AAA59486.1: JOINED.
 DR EMBL: M61935: AAA59486.1: JOINED.
 DR EMBL: M61936: AAA59486.1: JOINED.
 DR EMBL: M61938: AAA59486.1: JOINED.
 DR EMBL: M61939: AAA59486.1: JOINED.
 DR EMBL: M61940: AAA59486.1: JOINED.
 DR EMBL: M61941: AAA59486.1: JOINED.
 DR EMBL: M61942: AAA59486.1: JOINED.
 DR EMBL: M61943: AAA59486.1: JOINED.
 DR EMBL: M61944: AAA59486.1: JOINED.
 DR EMBL: M61945: AAA59486.1: JOINED.
 DR EMBL: M61947: AAA59486.1: JOINED.
 DR EMBL: M61948: AAA59486.1: JOINED.
 DR EMBL: M61949: AAA59486.1: JOINED.
 DR EMBL: M61950: AAA59486.1: JOINED.
 DR EMBL: M55370: AAA59485.1: JOINED.
 DR EMBL: M55378: AAA59485.1: JOINED.
 DR EMBL: M55365: AAA59485.1: JOINED.
 DR EMBL: M55371: AAA59485.1: JOINED.
 DR EMBL: M55372: AAA59485.1: JOINED.
 DR EMBL: M55373: AAA59485.1: JOINED.
 DR EMBL: M55374: AAA59485.1: JOINED.
 DR EMBL: M55375: AAA59485.1: JOINED.
 DR EMBL: M55376: AAA59485.1: JOINED.
 DR EMBL: M55344: AAA59485.1: JOINED.
 DR EMBL: M55345: AAA59485.1: JOINED.
 DR EMBL: M55347: AAA59485.1: JOINED.
 DR EMBL: M55348: AAA59485.1: JOINED.
 DR EMBL: M55349: AAA59485.1: JOINED.
 DR EMBL: M55350: AAA59485.1: JOINED.
 DR EMBL: M55351: AAA59485.1: JOINED.
 DR EMBL: M55352: AAA59485.1: JOINED.
 DR EMBL: M55353: AAA59485.1: JOINED.
 DR EMBL: M55355: AAA59485.1: JOINED.
 DR EMBL: M55356: AAA59485.1: JOINED.
 DR EMBL: M55357: AAA59485.1: JOINED.
 DR EMBL: M55358: AAA59485.1: JOINED.
 DR EMBL: M55359: AAA59485.1: JOINED.
 DR EMBL: M55360: AAA59485.1: JOINED.
 DR EMBL: M55361: AAA59485.1: JOINED.
 DR EMBL: M55362: AAA59485.1: JOINED.
 DR EMBL: M55363: AAA59485.1: JOINED.
 DR EMBL: M55364: AAA59485.1: JOINED.
 DR EMBL: M55366: AAA59485.1: JOINED.
 DR EMBL: M55367: AAA59485.1: JOINED.
 DR EMBL: M55368: AAA59485.1: JOINED.
 DR EMBL: M55369: AAA59485.1: JOINED.
 DR EMBL: M61916: AAA59482.1: -.
 DR EMBL: M20206: AAA59487.1: -.
 DR PIR: S13547: MMHUB1.
 DR HSP: P02468: 1KLO.
 DR Genew: HGNC:6486: LAMB1.
 DR MIM: 150240: -.
 DR Interpro: IPR000561: EGF-like.
 DR Interpro: IPR001886: Lamnt.
 DR Interpro: IPR002049: Laminin_EGF.
 DR Pfam: PF00053: Laminin_EGF_13.
 DR Pfam: PF00055: Laminin_Nterm_1.
 DR PRINTS: PRO0011: EGF_LAMININ.
 DR ProDom: PD002082: Lamnt: 1.
 DR SMART: SM00180: EGF_Lam: 11.
 DR SMART: SM00136: Lamnt: 1.
 DR PROSITE: PS00022: EGF_1: 9.
 DR PROSITE: PS01186: EGF_2: 2.
 DR PROSITE: PS01248: LAMININ_TYPE_EGF_11.
 DR GlycoProtex: Basement membrane: Extracellular matrix: Coiled coil;
 KW Laminin EGF-like domain: Cell adhesion: Repeat: Signal.
 FT SIGNL: 1 21
 FT CHAIN: 22 1786 LAMININ BETA-1 CHAIN.

FT DOMAIN 22 270 LAMININ N-TERMINAL (DOMAIN V1).
 FT DOMAIN 271 334 LAMININ EGF-LIKE 1.
 FT DOMAIN 335 397 LAMININ EGF-LIKE 2.
 FT DOMAIN 398 437 LAMININ EGF-LIKE 3.
 FT DOMAIN 438 509 LAMININ EGF-LIKE 4.
 FT DOMAIN 510 540 LAMININ EGF-LIKE 5 (INCOMPLETE).
 FT DOMAIN 541 771 LAMININ DOMAIN IV.
 FT DOMAIN 773 820 LAMININ EGF-LIKE 6.
 FT DOMAIN 821 866 LAMININ EGF-LIKE 7.
 FT DOMAIN 867 916 LAMININ EGF-LIKE 8.
 FT DOMAIN 917 975 LAMININ EGF-LIKE 9.
 FT DOMAIN 976 1027 LAMININ EGF-LIKE 10.
 FT DOMAIN 1028 1083 LAMININ EGF-LIKE 11.
 FT DOMAIN 1084 1131 LAMININ EGF-LIKE 12.
 FT DOMAIN 1132 1178 LAMININ EGF-LIKE 13.
 FT DOMAIN 1179 1397 DOMAIN II.
 FT DOMAIN 1398 1430 DOMAIN I.
 FT DOMAIN 1431 1786 DOMAIN I ALPHA.
 FT DOMAIN 1216 1315 COILED COIL (POTENTIAL).
 FT DOMAIN 1316 1353 COILED COIL (POTENTIAL).
 FT DOMAIN 1354 1388 COILED COIL (POTENTIAL).
 FT DOMAIN 1442 1781 BY SIMILARITY.
 FT DISULFID 271 280 BY SIMILARITY.
 FT DISULFID 273 298 BY SIMILARITY.
 FT DISULFID 300 309 BY SIMILARITY.
 FT DISULFID 312 332 BY SIMILARITY.
 FT DISULFID 335 344 BY SIMILARITY.
 FT DISULFID 337 362 BY SIMILARITY.
 FT DISULFID 365 374 BY SIMILARITY.
 FT DISULFID 377 395 BY SIMILARITY.
 FT DISULFID 398 411 BY SIMILARITY.
 FT DISULFID 400 426 BY SIMILARITY.
 FT DISULFID 428 437 BY SIMILARITY.
 FT DISULFID 440 455 BY SIMILARITY.
 FT DISULFID 448 472 BY SIMILARITY.
 FT DISULFID 460 479 BY SIMILARITY.
 FT DISULFID 481 490 BY SIMILARITY.
 FT DISULFID 493 507 BY SIMILARITY.
 FT DISULFID 773 785 BY SIMILARITY.
 FT DISULFID 775 792 BY SIMILARITY.
 FT DISULFID 794 803 BY SIMILARITY.
 FT DISULFID 806 818 BY SIMILARITY.
 FT DISULFID 821 833 BY SIMILARITY.
 FT DISULFID 823 840 BY SIMILARITY.
 FT DISULFID 842 851 BY SIMILARITY.
 FT DISULFID 854 864 BY SIMILARITY.
 FT DISULFID 867 876 BY SIMILARITY.
 FT DISULFID 869 883 BY SIMILARITY.
 FT DISULFID 886 895 BY SIMILARITY.
 FT DISULFID 898 914 BY SIMILARITY.
 FT DISULFID 917 933 BY SIMILARITY.

Query Match 51.3%. Score 40: DB 1; Length 1786;
 Best Local Similarity 61.5%; Pred. No. 1.2e+02;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 LNPNYLTKSDVT 14
 1111:1111
 DB 89 LNPDSHLEENVY 101

RESULT 10
 ID LMB1_MOUSE STANDARD: PRT: 1786 AA.
 AC P02469:
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Laminin beta-1 chain precursor (Laminin BI chain).
 GN LAMB1-1 OR LAMB-1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090.

[1]
RP SEQUENCE FROM N.A.
RA MEDLINE=87147212; PubMed=3493487;
RA Sasaki M., Kato S., Kohno K., Martin G.R., Yamada Y.;
RT "Sequence of the cDNA encoding the laminin B1 chain reveals a
RT multidomain protein containing cysteine-rich repeats";
RL Proc. Natl. Acad. Sci. U.S.A. 84:935-939(1987).
RN [2]
RP SEQUENCE OF 1292-1786 FROM N.A.
RA MEDLINE=85051302; PubMed=6209134;
RA Barlow D.P., Green N.M., Kurkinen M., Hogan B.L.M.;
RT "Sequencing of laminin B chain cDNAs reveals C-terminal regions of
RT coiled-coil alpha-helix";
RL EMBO J. 3:2355-2362(1984).
RN [3]
RP SEQUENCE OF 165-172; 539-547 AND 712-719.
RA STRAIN-BALB/C; TISSUE=Endothelial cells;
RA MEDLINE=97363207; PubMed=9219532;
RA Frieser M., Noeckel H., Pausch F., Roeder C., Hahn A., Deutzmann R.,
RA Sorokin L.M.;
RT "Cloning of the mouse laminin alpha 4 cDNA. Expression in a subset of
RT endothelium";
RL Eur. J. Biochem. 246:727-735(1997).
CC -1- FUNCTION: Binding to cells via a high affinity receptor, laminin
CC is thought to mediate the attachment, migration, and organization
CC of cells into tissues during embryonic development by interacting
CC with other extracellular matrix components.
CC -1- SUBUNIT: Laminin is a complex glycoprotein, consisting of three
CC different polypeptide chains (alpha, beta, gamma), which are bound
CC to each other by disulfide bonds into a cross-shaped molecule
CC comprising one long and three short arms with globules at each
CC end.
CC THE BETA-1 CHAIN IS A SUBUNIT OF LAMININ-1 (EHS LAMININ), LAMININ-
CC 2 (MEROSIN), AND LAMININ-6 (K-LAMININ).
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- TISSUE SPECIFICITY: FOUND IN THE BASEMENT MEMBRANES (MAJOR
CC COMPONENT).
CC -1- SIMILARITY: CONTAINS 1 LAMININ N-TERMINAL DOMAIN (DOMAIN VI).
CC -1- SIMILARITY: CONTAINS 12.5 LAMININ EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 LAMININ DOMAIN IV.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: M15525; AAA39407.1; ALT_INIT.
DR EMBL: X05212; CAA28839.1; -.
DR PIR: A26413; MMSB1..
DR HSSP: P02468; IKLO.
DR MGD: MGI:96743; Lambl-1.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001886; LAMNT.
DR InterPro: IPR002049; Laminin_EGF.
DR Pfam: PF00053; Laminin_EGF_13.
DR Pfam: PF00055; Laminin_Nterm; 1.
DR PRINTS: PR00011; EGF_LAMININ.
DR PRODOM: PD002082; LAMNT; 1.
DR SMART: SM00180; EGF_Lam; 11.
DR SMART: SM00136; LAMNT; 1.
DR PROSITE: PS00022; EGF_1; 9.
DR PROSITE: PS01186; EGF_2; 2.
DR PROSITE: PS01248; LAMININ_Type_EGF; 11.
DR GlycoProtein: Basement membrane; Extracellular matrix; Coiled coil;
DR Laminin EGF-like domain; Cell adhesion; Repeat; Signal.
KW SIGNAL
FT CHAIN 1 21
FT CHAIN 22 1786 LAMININ BETA-1 CHAIN.
FT DOMAIN 22 270 LAMININ N-TERMINAL (DOMAIN VI).
FT DOMAIN 271 334 LAMININ EGF-LIKE 1.
FT DOMAIN 335 397 LAMININ EGF-LIKE 2.

FT DOMAIN 398 457 LAMININ EGF-LIKE 3.
FT DOMAIN 458 509 LAMININ EGF-LIKE 4.
FT DOMAIN 510 540 LAMININ EGF-LIKE 5 (INCOMPLETE).
FT DOMAIN 541 772 LAMININ DOMAIN IV.
FT DOMAIN 773 820 LAMININ EGF-LIKE 6.
FT DOMAIN 821 866 LAMININ EGF-LIKE 7.
FT DOMAIN 867 916 LAMININ EGF-LIKE 8.
FT DOMAIN 917 975 LAMININ EGF-LIKE 9.
FT DOMAIN 976 1027 LAMININ EGF-LIKE 10.
FT DOMAIN 1028 1083 LAMININ EGF-LIKE 11.
FT DOMAIN 1084 1131 LAMININ EGF-LIKE 12.
FT DOMAIN 1132 1178 LAMININ EGF-LIKE 13.
FT DOMAIN 1179 1397 DOMAIN 11.
FT DOMAIN 1398 1430 DOMAIN ALPHA.
FT DOMAIN 1431 1786 DOMAIN I.
FT DOMAIN 1787 1882 COILED COIL (POTENTIAL).
FT DOMAIN 1883 1988 COILED COIL (POTENTIAL).
FT DOMAIN 1989 2271 COILED COIL (POTENTIAL).
FT DOMAIN 2272 2773 BY SIMILARITY.
FT DOMAIN 2774 3009 BY SIMILARITY.
FT DOMAIN 3010 3332 BY SIMILARITY.
FT DOMAIN 3333 3444 BY SIMILARITY.
FT DOMAIN 3445 3622 BY SIMILARITY.
FT DOMAIN 3623 3743 BY SIMILARITY.
FT DOMAIN 3744 3959 BY SIMILARITY.
FT DOMAIN 3960 4111 BY SIMILARITY.
FT DOMAIN 4112 4266 BY SIMILARITY.
FT DOMAIN 4267 4426 BY SIMILARITY.
FT DOMAIN 4427 4555 BY SIMILARITY.
FT DOMAIN 4556 4722 BY SIMILARITY.
FT DOMAIN 4723 4907 BY SIMILARITY.
FT DOMAIN 4908 5077 BY SIMILARITY.
FT DOMAIN 5078 5785 BY SIMILARITY.
FT DOMAIN 5786 7922 BY SIMILARITY.
FT DOMAIN 7923 8093 BY SIMILARITY.
FT DOMAIN 8094 8188 BY SIMILARITY.
FT DOMAIN 8189 8333 BY SIMILARITY.
FT DOMAIN 8334 8400 BY SIMILARITY.
FT DOMAIN 8401 8511 BY SIMILARITY.
FT DOMAIN 8512 8641 BY SIMILARITY.
FT DOMAIN 8642 8766 BY SIMILARITY.
FT DOMAIN 8767 8833 BY SIMILARITY.
FT DOMAIN 8834 8995 BY SIMILARITY.
FT DOMAIN 8996 9144 BY SIMILARITY.
FT DOMAIN 9145 9333 BY SIMILARITY.
FT DOMAIN 9334 9444 BY SIMILARITY.
FT DOMAIN 9445 9555 BY SIMILARITY.
FT DOMAIN 9556 9733 BY SIMILARITY.
FT DOMAIN 9734 9909 BY SIMILARITY.
FT DOMAIN 9910 10099 BY SIMILARITY.
FT DOMAIN 10100 10255 BY SIMILARITY.
FT DOMAIN 10256 10966 BY SIMILARITY.
FT DOMAIN 10967 11103 BY SIMILARITY.
FT DOMAIN 11104 11144 BY SIMILARITY.
FT DOMAIN 11145 11129 BY SIMILARITY.
FT DOMAIN 11130 11344 BY SIMILARITY.
FT DOMAIN 11345 11622 BY SIMILARITY.
FT DOMAIN 11623 11779 BY SIMILARITY.
FT DOMAIN 11780 11882 INTERCHAIN (PROBABLE).
FT DOMAIN 11883 1785 INTERCHAIN (PROBABLE).
FT DOMAIN 1786 1200 N-LINKED (GLCNAC).
FT DOMAIN 1201 120 N-LINKED (GLCNAC).
FT DOMAIN 1202 356 N-LINKED (GLCNAC).
FT DOMAIN 357 519 N-LINKED (GLCNAC).
FT DOMAIN 520 677 N-LINKED (GLCNAC).
FT DOMAIN 678 1041 N-LINKED (GLCNAC).
FT DOMAIN 1042 1195 N-LINKED (GLCNAC).
FT DOMAIN 1196 1279 N-LINKED (GLCNAC).
FT DOMAIN 1280 1336 N-LINKED (GLCNAC).

FT CARBOHYD 1343 1343 N-LINKED (GLICNAC. . .) (POTENTIAL).
FT CARBOHYD 1487 1487 N-LINKED (GLICNAC. . .) (POTENTIAL).
FT CARBOHYD 1533 1533 N-LINKED (GLICNAC. . .) (POTENTIAL).
FT CARBOHYD 1542 1542 N-LINKED (GLICNAC. . .) (POTENTIAL).
FT CARBOHYD 1643 1643 N-LINKED (GLICNAC. . .) (POTENTIAL).
FT CONFLICT 1531 1534 SGNA -> MEMP (IN REF. 2).
FT CONFLICT 1749 1749 D -> N (IN REF. 2).
SO SEQUENCE 1786 AA; 196904 MW; 846571B7BF41A474 CRC64;

Query Match 51.3%; Score 40; DB 1; Length 1786;
Best Local Similarity 61.5%; Pred. No. 1.2e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Db 2 LNPDNYLISKDYT 14
89 LNPDHSLIENVY 101

RESULT 11
PGKT_THEME STANDARD; PRT; 654 AA.
AC P36204; Q60031; 01-JUN-1994 (Rel. 29, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Bifunctional PCK/TIM (Includes: Phosphoglycerate kinase (EC 2.7.2.3);
DE Triosephosphate isomerase (EC 5.3.1.1) (TIM)).
GN PCK/TPI OR TM0689.
OS Thermotoga maritima.
OC Bacteria; Thermotogae; Thermotogales;
OC Thermotogaceae; Thermotoga.
OX NCBI_TaxID=2336;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=MSB8 / DSM 3109;
RX MEDLINE=95163577; PubMed=7859734;
RA Schurig H., Beaucamp N., Ostendorf R., Jaenicke R., Adler E.,
RA Knowles J.R.;
RT "Phosphoglycerate kinase and triosephosphate isomerase from the
RT hyperthermophilic bacterium Thermotoga maritima form a covalent
RT bifunctional enzyme complex.";
RL EMBO J. 14:442-451(1995).
[2]
RN SEQUENCE FROM N.A.
RP STRAIN=MSB8 / DSM 3109;
RX MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L., Utterback T.R., Malek J.A., Liner K.D., Garrett M.M.,
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips G.A., Richardson D.,
RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from
RT genome sequence of Thermotoga maritima.";
RL Nature 399:323-326(1999).
[3]
RN X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF PCK AND REVISIONS.
RP STRAIN=MSB8 / DSM 3109;
RX MEDLINE=98046096; PubMed=9384563;
RA Auerbach G., Huber R., Graetlinger M., Zaiss K., Schurig H.,
RA Jaenicke R., Jacob U.;
RT "Closed structure of phosphoglycerate kinase from Thermotoga maritima
RT reveals the catalytic mechanism and determinants of thermal
RT stability.";
RL Structure 5:1475-1483(1997).
[4]
RN X-RAY CRYSTALLOGRAPHY (2.85 ANGSTROMS) OF TIM.
RP MEDLINE=20058648; PubMed=10591103;
RX Maes D., zeelen J.P., Thanki N., Beaucamp N., Alvarez M., Thi M.H.,
RA Bachmann J., Martini J.A., Wynn L., Jaenicke R., Wierenga R.K.;
RT "The crystal structure of triosephosphate isomerase (TIM) from
RT Thermotoga maritima: a comparative thermostability structural
RT analysis of ten different TIM structures.";

RL Proteins 37:441-453(1999).
CC -1- CATALYTIC ACTIVITY: ATP + 3-phospho-D-glycerate = ADP + 3-
CC phospho-D-glyceroyl phosphate.
CC -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate = glycero-
CC phosphate.
CC -1- PATHWAY: Second phase of glycolysis; second step.
CC -1- SUBUNIT: MONOMER (PCK) AND HOMOTETRAMER (PCK-TIM).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE
CC PHOSPHOGYCLERATE KINASE FAMILY.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE
CC TRIOSEPHOSPHATE ISOMERASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: X75437; CAA53187.1; -
DR EMBL: L27492; AAB67520.1; -
DR EMBL: AE001741; AAD35771.1; -
DR PIR: S38439; S38439.
DR PDB: 1B9B; 01-JUN-00.
DR PDB: 1VPE; 17-JUN-98.
DR TIGR: TM0689; -
DR InterPro: IPR001576; PCK.
DR InterPro: IPR000652; Triophos_ismrse.
DR Pfam: PF001021; TIM; 1.
DR Pfam: PF00162; PCK; 1.
DR PRINTS: PR00477; PHGLICKINASE.
DR ProDom: PD001005; Triophos_ismrse; 1.
DR TIGRFAMs: TIGR00419; tim; 1.
DR PROSITE: PS00111; POLYMERASE_KINASE; 1.
DR PROSITE: PS00171; TIM; 1.
KW Transferase; Kinase; Isomerase; Glycolysis; Gluconeogenesis;
KW Fatty acid biosynthesis; Pentose shunt; Multifunctional enzyme;
KW 3D-structure; Complete proteome.
FT DOMAIN 1 399
FT DOMAIN 400 654
FT ACT_SITE 495
FT ACT_SITE 567
FT CONFLICT 213
FT CONFLICT 394
FT CONFLICT 626
FT CONFLICT 640
FT CONFLICT 640
SO SEQUENCE 654 AA; 71585 MW; 42358A4EF0C5E481 CRC64;

Query Match 50.6%; Score 39.5; DB 1; Length 654;
Best Local Similarity 38.9%; Pred. No. 49;
Matches 7; Conservative 5; Mismatches 3; Indels 3; Gaps 1;

Db 1 ALNPDNYLISKDYNG 15
613 SIKPDNGLITVQKIDG 630

RESULT 12
NUCA_ANASP
ID NUCA_ANASP STANDARD; PRT; 274 AA.
AC P38446;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Nuclease precursor (EC 3.1.30.-) (Endonuclease).
GN NUCA OR ALL7362.
OS Anabaena sp. (strain PCC 7120).
OG Plasmid pCC7120alpha.
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
[1]

RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RX MEDLINE-93125124; PubMed-1343821;
 RA Muro-Pastor A.M., Flores E., Herrero A., Molk C.P.;
 RT "Identification, genetic analysis and characterization of a
 RT sugar-non-specific nuclease from the cyanobacterium *Anabaena* sp. PCC
 RT 7120.";
 RL Mol. Microbiol. 6:3021-3030(1992).
 RM [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-21595285; PubMed-11759840;
 RA Kurek T., Nakamura Y., Molk C.P., Kunitz T., Sasamoto S.,
 RA Matsunaga A., Igarashi M., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kohara M., Matsunaga M., Matsuno A., Muraki A.,
 RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,
 RA Yasuda M., Tabata S.;
 RT "Complete genomic sequence of the filamentous nitrogen-fixing
 RT cyanobacterium *Anabaena* sp. strain PCC 7120.";
 RL DNA Res. 8:205-213(2001).
 CC -1- FUNCTION: CATALYZES THE DEGRADATION OF BOTH RNA AND DNA, HAS THE
 CC POTENTIAL TO ACT AS AN ENDONUCLEASE.
 CC -1- COPACITOR: REQUIRES A DIVALENT CATION FOR ACTIVITY; THE
 CC EFFECTIVENESS OF THE CATIONS ARE $Mg^{2+} > Mg^{2+} > Ca^{2+} = Co^{2+}$.
 CC -1- SUBCELLULAR LOCATION: PERIPLASMIC OR LOOSELY ATTACHED TO THE
 CC CYTOPLASMIC OR THE OUTER MEMBRANE.
 CC -1- PPM: THE N-TERMINUS IS BLOCKED.
 CC -1- SIMILARITY: BELONGS TO THE DNA/RNA NON-SPECIFIC ENDONUCLEASES
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X64705; CAA45962.1; -;
 DR EMBL: AP003601; BAB77120.1; -;
 DR PIR: S28039; S28039;
 DR InterPro: IPR001604; Endonuclease.
 DR Pfam: PF01223; Endonuclease; 1.
 DR SMART: SM00477; NUC; 1.
 DR PROSITE: PS01070; NUCLEASE_NON_SPEC; 1.
 KW Hydrolyase; Nuclease; Endonuclease; Manganese; Signal;
 KW Periplasmic; Plasmid; Complete proteome.
 FT SIGNAL 1 24 POTENTIAL.
 FT CHAIN 1 25 274 NUCLEASE.
 FT ACT_SITE 124 124 BY SIMILARITY.
 SO SEQUENCE 274 AA; 23669 MW; 7F14170915AE59BF CRC64;
 Query Match 50.0%; Score 39; DB 1; Length 274;
 Best Local Similarity 63.6%; Pred. No. 23;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 Oy 2 LNPNTYLSKID 12
 Db 53 LTPDNYLMVKN 63
 RESULT 13
 ROX1_YEAST STANDARD: PRT; 368 AA.
 ID ROX1_YEAST
 AC P25042;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE ROX1 repressor (Hypoxia function repressor) (Heme-dependent repression
 DE factor).
 GN ROX1 OR YPR065W OR YPR949.20.
 OS Saccharomyces cerevisiae (Baker's yeast).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomycaceae;
 CC NCBI_taxonomy:4932;

RM [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-94019282; PubMed-8413209;
 RA Balasubramanian B., Lowry C.V., Zitomer R.S.;
 RT "The rox1 repressor of the *Saccharomyces cerevisiae* hypoxic genes is
 RT a specific DNA-binding protein with a high-mobility-group motif.";
 RL Mol. Cell. Biol. 13:6071-6078(1993).
 RM [2]
 RP SEQUENCE FROM N.A.
 RX STRAIN-8288C / AB972;
 RA Bussey H., Storms R.K., Ahmed A., Albermann K., Allen E., Anstorge W.,
 RA Aurrejo R., Aparicio A., Barrell B.G., Badcock K., Benes V.,
 RA Boekestein D., Bowman S., Buckner M., Carpenter J., Cherry J.M.,
 RA Chung E., Churcher C.M., Coster F., Davis K., Davis R.W.,
 RA Dietrich F.S., Delius H., DiPaolo T., Dubois E., Duesterhoeft A.,
 RA Duncan M., Floeth M., Fortin N., Friesen J.D., Fritz C., Goffeau A.,
 RA Hall J., Hebling U., Heumann K., Hilbert H., Hillier L.,
 RA Hunnicke-Smith S., Hyman R., Johnston M., Kallman S., Klein R.,
 RA Komp C., Kurdi O., Lashkari D., Lew H., Lin A., Lin D., Louis E.J.,
 RA Marathe R., Messenguy F., Mewes H.-W., Mirzaii S., Moestl D.,
 RA Mueller-Ruer S., Nemeth A., Newbigh U., Oefner P., Pearson D.,
 RA Petel F.X., Pohl T.M., Purnelle D., Schaefer M., Scharte M.,
 RA Scherens B., Schramm S., Schroeder M., Sidou A.M., Tietelin H.,
 RA Uristarazu L.A., Ushinsky S., Vierdeels F., Vissers S., Voss H.,
 RA Walsh S.V., Wambut R., Wang Y., Wedler E., Wedler H., Winnett E.,
 RA Zhong W.W., Zollner A., Vo D.H., Han J.;
 RT "The nucleotide sequence of *Saccharomyces cerevisiae* chromosome XVI.";
 RL Nature 387:103-105(1997).
 RM [3]
 RP CHARACTERIZATION.
 RX MEDLINE-96174644; PubMed-8600445;
 RA di Flumeri C., Liston P., Acheson N.H., Keng T.;
 RT "The HMG domain of the ROX1 protein mediates repression of HEM13
 RT through overlapping DNA binding and oligomerization functions.";
 RL Nucleic Acids Res. 24:808-815(1996).
 CC -1- FUNCTION: TRANSCRIPTION FACTOR THAT REPRESSSES THE EXPRESSION OF
 CC HEM13, COX5B, ANB1, CYC7 OR AMC3 (HYPOXIC FUNCTION). BINDS TO THE
 CC DNA SEQUENCE 5'-RRRTAACAGAG-3'.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- INDUCTION: BY HEME.
 CC -1- SIMILARITY: CONTAINS 1 HMG BOX.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X60458; CAA42991.1; -;
 DR EMBL: 249219; CAA89182.1; -;
 DR EMBL: 271255; CAA94973.1; -;
 DR PIR: S17015; S17015;
 DR HSP: O05066; HMY;
 DR TRANSFAC: T01286;
 DR SGD: S0006269; ROX1;
 DR InterPro: IPR000910; HMG_12_box.
 DR Pfam: PF00505; HMG_box; 1.
 DR SMART: SM00398; HMG; 1.
 KW Transcription regulation; Repressor; DNA-binding; Nuclear protein.
 FT DNA_BIND 14 83 HMG BOX.
 FT DOMAIN 102 123 GLN-RICH.
 SO SEQUENCE 368 AA; 41838 MW; 3B27442D7DEE3DBD CRC64;
 Query Match 50.0%; Score 39; DB 1; Length 368;
 Best Local Similarity 33.3%; Pred. No. 32;
 Matches 5; Conservative 7; Mismatches 3; Indels 0; Gaps 0;
 Oy 1 ALNPNTYLSKIDVTG 15
 Db 165 SVNTSNVMSRSLSG 179

RESULT 14
CD36_HUMAN STANDARD: PRT: 471 AA.
ID P16671; Q16093; Q13966;
AC 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Platelet glycoprotein IV (GP1b) (GP1b) (CD36 antigen) (PAS IV)
DE (PAS 4 protein).
CN CD36 OR GP4 OR GP3B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=89324065; PubMed=2473841;
RA Oquendo P., Hundt E., Lawler J., Seed B.;
RT "CD36 directly mediates cytoadherence of Plasmodium falciparum
parasitized erythrocytes."
RL Cell 58:95-101(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC Sugimoto Y., Tsuruo T.;
RX Submitted (xxx-1992) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=94308159; PubMed=7518447;
RA Armegilla A.L., Vega M.A.;
RT "Structural organization of the gene for human CD36 glycoprotein."
RL J. Biol. Chem. 269:18985-18991(1994).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Platelet;
RX MEDLINE=94082337; PubMed=7505064;
RA Wyler B., Davlet L., Borkiewicz H., Bortel J.C., McGregor J.L.;
RT "Cloning of the cDNA encoding human platelet CD36: comparison to PCR
amplified fragments of monocyte, endothelial and HEL cells."
RL Thromb. Haemost. 70:500-505(1993).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=94040811; PubMed=7593552;
RA Taylor K.T., Tang Y., Sodisaki D.A., Lipsky R.H.;
RT "Characterization of two alternatively spliced 5'-untranslated exons
of the human CD36 gene in different cell types."
RL Gene 133:205-212(1993).
RN [6]
RP SEQUENCE OF 1-36.
RC TISSUE=Platelet;
RX MEDLINE=89214205; PubMed=2468669;
RA Tandon N.N., Lipsky R.H., Burgess W.H., Jamieson G.A.;
RT "Isolation and characterization of platelet glycoprotein IV (CD36)."
RL J. Biol. Chem. 264:7570-7575(1989).
RN [7]
RP PALMITOYLATION OF CYS-2, CYS-6, CYS-463 AND CYS-465.
RX MEDLINE=96394284; PubMed=8798390;
RA Tao N., Wagner S.J., Lublin D.M.;
RT "CD36 is palmitoylated on both N- and C-terminal cytoplasmic tails."
RL J. Biol. Chem. 271:22315-22320(1996).
RN [8]
RP FUNCTION: SEEMS TO HAVE NUMEROUS POTENTIAL PHYSIOLOGICAL
FUNCTIONS. BINDS TO COLLAGEN, THROMBOSPONDIN, ANIONIC
PHOSPHOLIPIDS AND OXIDIZED LDL. MAY FUNCTION AS A CELL ADHESION
MOLECULE. DIRECTLY MEDIATES CYTOADHERENCE OF PLASMODIUM FALCIPARUM
PARASITIZED ERYTHROCYTES. BINDS LONG CHAIN FATTY ACIDS AND MAY
FUNCTION IN THE TRANSPORT AND/OR AS A REGULATOR OF FATTY ACID
TRANSPORT.
RN [9]
RP SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO THE CD36 FAMILY.
CC -1- DATABASE: NAME=PROW; NOTE=CD guide CD36 entry;
WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd36.htm".

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: M24795; AAA35534.1;
DR EMBL: M98398; AAA58412.1;
DR EMBL: M98399; AAA58413.1;
DR EMBL: 232770; CAAB3662.1;
DR EMBL: 232754; CAAB3662.1; JOINED.
DR EMBL: 232755; CAAB3662.1; JOINED.
DR EMBL: 232756; CAAB3662.1; JOINED.
DR EMBL: 232757; CAAB3662.1; JOINED.
DR EMBL: 232758; CAAB3662.1; JOINED.
DR EMBL: 232759; CAAB3662.1; JOINED.
DR EMBL: 232760; CAAB3662.1; JOINED.
DR EMBL: 232761; CAAB3662.1; JOINED.
DR EMBL: 232762; CAAB3662.1; JOINED.
DR EMBL: 232763; CAAB3662.1; JOINED.
DR EMBL: 232764; CAAB3662.1; JOINED.
DR EMBL: S67532; AAD13993.1;
DR EMBL: L06850; AAA16068.1;
DR PIR: A30989; A30989.
DR PIR: A33357; A33357.
DR Gene: HGNC:1663; CD36.
DR InterPro: IPR002159; CD36.
DR Pfam: PF01130; CD36.1.
KW Glycoprotein; Transmembrane; Antigen; Cell adhesion; Transport;
KW Receptor; Lipoprotein; Palmitate.
FT INIT_MET 0
FT DOMAIN 1 6
FT TRANSMEM 7 28
FT DOMAIN 29 438
FT TRANSMEM 439 460
FT DOMAIN 461 471
FT LIPID 2 2
FT LIPID 6 6
FT LIPID 463 463
FT LIPID 465 465
FT DIOLFID 242 310
FT DIOLFID 271 332
FT DIOLFID 312 321
FT CARBOHYD 78 78
FT CARBOHYD 101 101
FT CARBOHYD 133 133
FT CARBOHYD 162 162
FT CARBOHYD 204 204
FT CARBOHYD 219 219
FT CARBOHYD 234 234
FT CARBOHYD 246 246
FT CARBOHYD 320 320
FT CARBOHYD 376 376
FT CONFLICT 43 43
FT CONFLICT 237 237
FT CONFLICT 373 373
SO SEQUENCE 471 AA; 52922 MW; 61B1B5E7D661F161 CRC64;

Query Match 50.08; Score 39; DB 1; Length 471;
Best Local Similarity 50.08; Pred. No. 42;
Matches 9; Conservative 2; Mismatches 1; Indels 6; Gaps 1;

OY 3 NPDPV-----LISKDVT 14
DB 305 NPDPVCTEKTISKNT 322

RESULT 15
LGUL_SCHPO

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd

OM protein - protein search, using SW model

Run on: June 26, 2003, 00:47:39 ; Search time 4.58824 Seconds
(without alignments)
673.616 Million cell updates/sec

Title:	US-09-980-370-1
Perfect score:	78
Sequence:	1 ALNPDNYLISKDVTG 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 08

Listing first 45 summaries

Database :

```

1:  sp.archaea:*
2:  sp.bacteria:*
3:  sp.fungi:*
4:  sp.human:*
5:  sp.invertebrate:*
6:  sp.mammal:*
7:  sp.mhc:*
8:  sp.organelle:*
9:  sp.phage:*
10: sp.plant:*
11: sp.potent:*
12: sp.virus:*
13: sp.vertebrate:*
14: sp.unclassified:*
15: sp.vlvs:*
16: sp.bacteriap:*
17: sp.archaea:*

```

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	78	100.0	1097	2	P72196	P72196 porphyromon
2	78	100.0	1223	2	Q9ZMB5	Q9ZMB5 porphyromon
3	78	100.0	1358	2	P69697	P69697 porphyromon
4	78	100.0	1687	2	Q9R987	Q9R987 porphyromon
5	78	100.0	1704	2	O51816	O51816 porphyromon
6	78	100.0	1706	2	O51839	O51839 porphyromon
7	78	100.0	1706	2	O51838	O51838 porphyromon
8	78	100.0	1723	2	P72194	P72194 porphyromon
9	78	100.0	1723	2	P72197	P72197 porphyromon
10	78	100.0	1732	2	O52050	O52050 porphyromon
11	78	100.0	1732	2	O51817	O51817 porphyromon
12	78	100.0	1734	2	O07442	O07442 porphyromon
13	47	50.3	134	5	O55965	O55965 caenorhabdi
14	46	59.0	1379	5	Q9BH56	Q9BH56 caenorhabdi
15	45	57.7	312	4	Q9K1B3	Q9K1B3 porphyromon
16	45	57.7	335	10	Q8W4Z0	Q8W4Z0 medicago sc

ALIGNMENTS

17	45	57.7	695	11	054767	054767 rattus novae
18	45	57.7	841	10	09C552	09C552 arbidopsis
19	44	56.4	315	10	09FX92	09FX92 arbidopsis
20	44	56.4	499	10	09JY34	09JY34 arbidopsis
21	44	56.4	530	10	09PHJ3	09PHJ3 arbidopsis
22	44	56.4	630	12	09BH1	09BH1 emsacta moos
23	44	56.4	639	16	09CES3	09CES3 lactococcus
24	44	56.4	752	17	09H5A1	09H5A1 halobacter
25	43	55.1	121	8	09T1W	09T1W cyanidium c
26	43	55.1	120	10	04B69	04B69 arbidopsis
27	43	55.1	514	5	09VX24	09VX24 arbidopsis
28	43	55.1	925	2	09R10	09R10 porphyromon
29	43	55.1	2877	12	05C678	05C678 peanut bud
30	43	55.1	2878	12	09NS01	09NS01 watermelon
31	43	55.1	318	10	09LMA7	09LMA7 arbidopsis
32	42	53.8	374	10	09FX53	09FX53 arbidopsis
33	42	53.8	500	10	094D20	094D20 oryza sativi
34	42	53.8	674	5	022578	022578 caenorhabdi
35	42	53.8	1844	5	022579	022579 caenorhabdi
36	42	53.8	2548	4	09UMJ2	09UMJ2 homo sapien
37	42	53.8	2165	17	08B169	08B169 melanosarc
38	41.5	53.2	146	10	08B391	08B391 mesembryan
39	41	52.6	256	11	09P659	09P659 mus musculi
40	41	52.6	369	16	09C4M8	09C4M8 pasteurirell
41	41	52.6	383	2	046190	046190 clostridium
42	41	52.6	397	5	09G052	09G052 thiospirillum
43	41	52.6	470	5	09VW6	09VW6 thiospirillum
44	41	52.6	485	2	045771	045771 bacteroides

RESULT 1
P73706

ID	PRELIMINARY:	PRT:	1097 AA.
AC	P72196.		
DT	01-FEB-1997 (TREMBLrel. 02, Created)		
DT	01-FEB-1997 (TREMBLrel. 02, Last sequence update)		
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)		
DE	Tomb-linked adhesin precursor.		
GN	TLA.		
OS	Porphyromonas gingivalis (Bacteroides gingivalis).		
OC	Bacteriia, CFb group: Bacteroidetes: Bacteroidales: Porphyromonadaceae:		
OC	Porphyromonas.		
OX	NCBI_TaxId=837:		
RN	11)		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=97386416; PubMed=9244265;		
RA	Aduse-Opoku J., Slaney J.M., Young K.A., Muir J., Rangarajan M.,		
RA	Curtis M.A.;		
RT	"The tla gene of Porphyromonas gingivalis M50: a homologue of the		
RT	arginine-specific protease precursor (PrPRI) which shares sequence		
RL	similarity to Tomb-linked receptors."		
RL	J. Bacteriol. 179:4778-4788(1997)."		
DR	EMBL: Y07618; CA668897.1; "		
DR	InterPro: IPR0009377; DNA_Ligase_C25.		
DR	InterPro: IPR001769; peptidase_C25.		
DR	Pfam: PF01364; Peptidase_C25; 3.		
DR	ProSite: PS00697; DNA_LIGASE_A1; UNKNOWN_1.		
KW	Signal.		
KT	SIGNAL		
SO	SEQUENCE	1	POTENTIAL.
		53	
	1097 AA:	118731 MW:	73BBA37B421PB89 CRC64:
Query Match		100.08:	Score 78: DB 2: Length 1097:
Best Local Similarity		100.08:	Pred. No. 8:1e-05:
Matches 15:	Conservative	0:	Mismatches 0:
			Indels 0:
			Gaps 0:
QV	1 ALMPDNYLISKDVGTG 15		
DB	581 ALMPDNYLISKDVGTG 595		

RESULT 2

09ZNB5 PRELIMINARY: PRT: 1223 AA.
 ID 09ZNB5
 AC 09ZNB5
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE 130k-HMGD (fragment).
 OS Porphyromonas gingivalis (Bacteroides gingivalis).
 OC Bacteroides; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
 OC Porphyromonas group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
 OC NCBI_TaxID=837;
 CX NCBI_TaxID=837;
 RN 11
 RP SEQUENCE FROM N.A.
 RC STRAIN=381;
 RX MEDLINE=99143166; PubMed=9988746;
 RA Shibata Y., Hayakawa M., Takiguchi H., Shiroza T., Abiko Y.;
 RT "Determination and characterization of the hemagglutinin-associated
 RT short motifs found in Porphyromonas gingivalis multiple gene
 RT products."
 RL J. Biol. Chem. 274:5012-5020(1999).
 DR EMBL: AB019363; BAA34341.1;
 DR InterPro: IPR000977; DNA_Ligase.
 DR InterPro: IPR001769; Peptidase_C25.
 DR Pfam: PF01364; Peptidase_C25; 3.
 DR PROSITE: PS00697; DNA_Ligase_A1; UNKNOWN_1.
 FT NON_TER 1
 SQ SEQUENCE 1223 AA; 131542 MW; 00225CD2BA9P91B3 CRC64;

Query Match

Best Local Similarity 100.0%; Score 78; DB 2; Length 1223;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALPNPVLISKDVYG 15
 Db 707 ALPNPVLISKDVYG 721

RESULT 3

P96967 PRELIMINARY: PRT: 1358 AA.
 ID P96967
 AC P96967
 DT 01-MAY-1997 (TREMBLrel. 03, Created)
 DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Hemagglutinin.
 OS Porphyromonas gingivalis (Bacteroides gingivalis).
 OC Bacteroides; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
 OC Porphyromonas group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
 OC NCBI_TaxID=837;
 CX NCBI_TaxID=837;
 RN 11
 RP SEQUENCE FROM N.A.
 RC STRAIN=381;
 RA Han N., Lepine G., Whitlock J., Wojciechowski L., Progniske-Fox A.;
 RT "Cloning, sequencing and characterization of hagd, a member of the
 RT harsp multigene family in Porphyromonas gingivalis."
 RT Submitted (Aug-1996) to the EMBL/GenBank/DBJ databases.
 RL EMBL: U68489; A849691.1;
 DR InterPro: IPR001769; Peptidase_C25.
 DR InterPro: IPR001769; Peptidase_C25.
 DR Pfam: PF01364; Peptidase_C25; 3.
 DR PROSITE: PS00697; DNA_Ligase_A1; UNKNOWN_1.
 FT NON_TER 1
 SQ SEQUENCE 1358 AA; 147102 MW; 47FCA0B25B06DE08 CRC64;

Query Match

Best Local Similarity 100.0%; Score 78; DB 2; Length 1358;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALPNPVLISKDVYG 15
 Db 850 ALPNPVLISKDVYG 864

RESULT 4

09R9B7 PRELIMINARY: PRT: 1687 AA.
 ID 09R9B7
 AC 09R9B7
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Hemagglutinin/protease.
 OS Porphyromonas gingivalis (Bacteroides gingivalis).
 OC Bacteroides; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
 OC Porphyromonas group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
 OC NCBI_TaxID=837;
 CX NCBI_TaxID=837;
 RN 11
 RP SEQUENCE FROM N.A.
 RC STRAIN=381;
 RA Han N., Dong H., Progniske-Fox A.;
 RT "Cloning and characterization of hage from P. gingivalis 381."
 RT Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
 RL EMBL: AF026946; AAD01810.1;
 DR HSP: P95493; ICVR.
 DR MEROPS: C25.001;
 DR InterPro: IPR000977; DNA_Ligase.
 DR InterPro: IPR001769; Peptidase_C25.
 DR Pfam: PF01364; Peptidase_C25; 3.
 DR PROSITE: PS00697; DNA_Ligase_A1; UNKNOWN_1.
 FT PROTEASE 1
 SQ SEQUENCE 1687 AA; 183702 MW; D085B516A399FE70 CRC64;

Query Match

Best Local Similarity 100.0%; Score 78; DB 2; Length 1687;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALPNPVLISKDVYG 15
 Db 1179 ALPNPVLISKDVYG 1193

RESULT 5

051816 PRELIMINARY: PRT: 1704 AA.
 ID 051816
 AC 051816
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Arg-gingipain-1 proteinase.
 OS Porphyromonas gingivalis (Bacteroides gingivalis).
 OC Bacteroides; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
 OC Porphyromonas group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
 OC NCBI_TaxID=837;
 CX NCBI_TaxID=837;
 RN 11
 RP SEQUENCE FROM N.A.
 RC MEDLINE=95138080; PubMed=7836351;
 RA Ravidoff N., Potempa J., Pike R.N., Prochazka V., Kiefer M.C.,
 RA Travis J., Barr P.J.;
 RT "Molecular cloning and structural characterization of the Arg-
 RT gingipain proteinase of Porphyromonas gingivalis. Biosynthesis as a
 RT proteinase-adhesin polypeptide."
 RL J. Biol. Chem. 270:1007-1010(1995).
 DR EMBL: U15282; A469539.1;
 DR HSP: P95493; ICVR.
 DR MEROPS: C25.001;
 DR InterPro: IPR000977; DNA_Ligase.
 DR InterPro: IPR001769; Peptidase_C25.
 DR Pfam: PF01364; Peptidase_C25; 3.
 DR PROSITE: PS00697; DNA_Ligase_A1; UNKNOWN_1.
 FT CHAIN 228 719 MATURE 50-KDA CYSTEINE PROTEINASE
 FT CHAIN 228 719 MATURE 50-KDA CYSTEINE PROTEINASE
 SQ SEQUENCE 1704 AA; 185436 MW; 6A34B4013C2A676 CRC64;

Query Match

Best Local Similarity 100.0%; Score 78; DB 2; Length 1704;

Best Local Similarity 100.0%; Pred. No. 0.00013;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALNDPNYLISKDVTG 15
1196 ALNDPNYLISKDVTG 1210

RESULT 6
051839 PRELIMINARY: PRT: 1706 AA.

ID 051839
AC 051839: 051840;
DT 01-NOV-1996 (TRENBLREL. 01, Created)
DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
DT 01-JUN-2002 (TRENBLREL. 21, Last annotation update)
DE Arginine-specific thiol protease precursor.
GN PRT.
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; CFM group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
OC Porphyromonas.
OX NCBI_TaxID=837;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN-W50;
RX MEDLINE=95160709; PubMed=7857299;
RA Kirszenbaum L., Sotiropoulos C., Jackson C., Cleal S., Slakeski N.,
RA Reynolds E.C.;
RT "Complete nucleotide sequence of a gene prtr of Porphyromonas
gingivalis W50 encoding a 132 kDa protein that contains an arginine-
specific thiol endopeptidase domain and a haemagglutinin domain.";
RL Blochem. Biophys. Res. Commun. 207:424-431(1995).
[2]
RN SEQUENCE FROM N.A.
RP STRAIN-W50;
RC MEDLINE=96311339; PubMed=8713096;
RA Slakeski N., Cleal S.M., Reynolds E.C.;
RT "Characterization of a Porphyromonas gingivalis gene prtr that encodes
an arginine-specific thiol proteinase and multiple adhesins.";
RL Blochem. Biophys. Res. Commun. 224:605-610(1996).
[3]
RN SEQUENCE FROM N.A.
RP STRAIN-W50;
RC STRAIN-W50;
RA Reynolds E.;
RT Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.
[4]
RN SEQUENCE FROM N.A.
RP STRAIN-W50;
RC STRAIN-W50;
RA Slakeski N.;
RT Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL: L26341; AAC18876.1; -.
DR HSSP: P95493; 1CVR.
DR InterPro: IPR000977; DNA_ligase.
DR InterPro: IPR001769; Peptidase_C25.
DR Pfam: PF01364; Peptidase_C25; 3.
DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
GN KGP.
FT SIGNAL
FT CHAIN 1 23 POTENTIAL.
FT CHAIN 228 1706 ARGININE-SPECIFIC THIOI. PROTEASE.
SQ SEQUENCE 1706 AA; 185626 MW; E8BDFO7C9813844 CRC64;

Query Match 100.0%; Score 78; DB 2; Length 1706;
Best Local Similarity 100.0%; Pred. No. 0.00013;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALNDPNYLISKDVTG 15
1198 ALNDPNYLISKDVTG 1212

RESULT 7
051838 PRELIMINARY: PRT: 1706 AA.
ID 051838
AC 051838;

DT 01-NOV-1996 (TRENBLREL. 01, Created)
DT 01-MAY-1997 (TRENBLREL. 03, Last sequence update)
DT 01-JUN-2002 (TRENBLREL. 21, Last annotation update)
DE Protease precursor.
GN PRT.
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; CFM group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
OC Porphyromonas.
OX NCBI_TaxID=837;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN-W50;
RX MEDLINE=96071894; PubMed=7591131;
RA Aduse-Opoku J., Muir J., Slaney J.M., Rangarajan M., Curtis M.A.;
RT "Characterization, genetic analysis, and expression of a protease
antigen (PRT) of Porphyromonas gingivalis W50.";
RL Infect. Immun. 63:4744-4754(1995).
[2]
RN SEQUENCE FROM N.A.
RP STRAIN-W50;
RC Rangarajan M., Aduse-Opoku J., Slaney J.M., Young K.A., Curtis M.A.;
RT "The prtr1 and the prtr2 arginine-specific protease genes of
Porphyromonas gingivalis W50 produce five biochemically distinct
enzymes.";
RL Mol. Microbiol. 23:0-0(1997).
DR EMBL: X82680; CAA57997.1; -.
DR HSSP: P95493; 1CVR.
DR MEROPS: C25.001; -.
DR InterPro: IPR000977; DNA_ligase.
DR InterPro: IPR001769; Peptidase_C25.
DR Pfam: PF01364; Peptidase_C25; 3.
DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
GN SIGNAL
FT SIGNAL 1 23 POTENTIAL.
FT CHAIN 228 719 ALPHA-PROTEASE.
FT CHAIN 720 1262 BETA-ADHESIN.
SQ SEQUENCE 1706 AA; 185705 MW; 0E56DCD87FDA8CDD CRC64;

Query Match 100.0%; Score 78; DB 2; Length 1706;
Best Local Similarity 100.0%; Pred. No. 0.00013;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALNDPNYLISKDVTG 15
1198 ALNDPNYLISKDVTG 1212

RESULT 8
ID P72194
AC P72194: PRELIMINARY: PRT: 1723 AA.
DT 01-FEB-1997 (TRENBLREL. 02, Created)
DT 01-FEB-1997 (TRENBLREL. 02, Last sequence update)
DT 01-JUN-2002 (TRENBLREL. 21, Last annotation update)
DE Lys-gingipain.
GN KGP.
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; CFM group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
OC Porphyromonas.
OX NCBI_TaxID=837;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN-W50;
RC MEDLINE=97044756; PubMed=8889827;
RX Okamoto K., Kadowaki T., Nakayama K., Yamamoto K.;
RT "Cloning and sequencing of the gene encoding a novel lysine-specific
cysteine proteinase (Lys-gingipain) in Porphyromonas gingivalis.
RT structural relationship with the arginine-specific cysteine proteinase
(Arg-gingipain).";
RL J. Biochem. 120:398-406(1996).
DR EMBL: D83258; BAA11870.1; -.
DR MEROPS: C25.002; -.
DR InterPro: IPR000977; DNA_ligase.

DR InterPro: IPR001769; Peptidase_C25.
 DR Pfam: PF01364; Peptidase_C25; 3.
 DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
 SQ SEQUENCE 1723 AA; 187261 MW; 5628963D51493EB CRC64;

Query Match 100.0%; Score 78; DB 2; Length 1723;
 Best Local Similarity 100.0%; Pred. No. 0.00013;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNPDNYLISKDVTG 15
 |||
 DB 1215 ALNPDNYLISKDVTG 1229

RESULT 9

P72197 PRELIMINARY; PRT: 1723 AA.

AC P72197;
 DT 01-FEB-1997 (TREMBLrel. 02, Created)
 DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Lys-gingipain.
 GN KCP.
 OS Porphyromonas gingivalis (Bacteroides gingivalis).
 OC Bacteria; CFB group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
 OC Porphyromonas.
 NCBI_TaxID=837;
 RN [1]
 RA SEQUENCE FROM N.A.
 RA Pavloff N., Pemberton P.A., Potempa J., Chen W.-C.A., Pike R.N.,
 RA Prochazka V., Klefer M.C., Travis J., Barr P.J.;
 RC "Molecular cloning and characterization of Porphyromonas gingivalis
 RT Lys-gingipain.";
 RT Submitted (Apr. 1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U04691; AAA9810.1; -.
 DR MEROPS: C25.002; -.
 DR InterPro: IPR00977; DNA_Ligase.
 DR InterPro: IPR001769; Peptidase_C25.
 DR Pfam: PF01364; Peptidase_C25; 3.
 DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
 SQ SEQUENCE 1723 AA; 186831 MW; 4508A7E50197CEBD CRC64;

Query Match 100.0%; Score 78; DB 2; Length 1723;
 Best Local Similarity 100.0%; Pred. No. 0.00013;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNPDNYLISKDVTG 15
 |||
 DB 1215 ALNPDNYLISKDVTG 1229

RESULT 10

O52050 PRELIMINARY; PRT: 1732 AA.

DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Lysine specific cysteine protease.
 OS Porphyromonas gingivalis (Bacteroides gingivalis).
 OC Bacteria; CFB group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
 OC Porphyromonas.
 NCBI_TaxID=837;
 RN [1]
 RA SEQUENCE FROM N.A.
 RC STRAIN-W83.
 RX MEDLINE=98298016; PubMed=9632563;
 RA Lewis J.P., Macrina F.L.;
 RT "IS195, an insertion sequence-like element associated with protease
 genes in Porphyromonas gingivalis.";
 RL Infect. Immun. 66:3035-3042(1998).
 DR EMBL: AF017059; AAC26523.1; -.
 DR MEROPS: C25.002; -.

DR InterPro: IPR009977; DNA_Ligase.
 DR InterPro: IPR001769; Peptidase_C25.
 DR Pfam: PF01364; Peptidase_C25; 3.
 DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
 SQ SEQUENCE 1732 AA; 187931 MW; 82337463D5CB5EA5 CRC64;

Query Match 100.0%; Score 78; DB 2; Length 1732;
 Best Local Similarity 100.0%; Pred. No. 0.00013;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNPDNYLISKDVTG 15
 |||
 DB 1216 ALNPDNYLISKDVTG 1230

RESULT 11

O51817 PRELIMINARY; PRT: 1732 AA.

AC O51817;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Porphylin.
 GN PRP.
 OS Porphyromonas gingivalis (Bacteroides gingivalis).
 OC Bacteria; CFB group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
 OC Porphyromonas.
 NCBI_TaxID=837;
 RN [1]
 RA SEQUENCE FROM N.A.
 RC STRAIN-W12;
 RX MEDLINE=96213011; PubMed=8631659;
 RA Barkocy-Gallagher G.A., Han N., Patli J.M., Whitlock J.,
 RA Progulski-Fox A., Lantz M.S.;
 RC "Analysis of the prt gene encoding porphylin, a cysteine proteinase
 RT of Porphyromonas gingivalis.";
 RT J. Bacteriol. 178:2734-2741(1996).
 DR EMBL: U42210; AAB05655.1; -.
 DR MEROPS: C25.002; -.
 DR InterPro: IPR009977; DNA_Ligase.
 DR InterPro: IPR001769; Peptidase_C25.
 DR Pfam: PF01364; Peptidase_C25; 3.
 DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
 SQ SEQUENCE 1732 AA; 187875 MW; 654271DBE7BCAE4 CRC64;

Query Match 100.0%; Score 78; DB 2; Length 1732;
 Best Local Similarity 100.0%; Pred. No. 0.00013;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNPDNYLISKDVTG 15
 |||
 DB 1216 ALNPDNYLISKDVTG 1230

RESULT 12

O07442 PRELIMINARY; PRT: 1732 AA.

AC O07442;
 DT 01-JUL-1997 (TREMBLrel. 04, Created)
 DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Lysine-specific cysteine proteinase.
 GN PRP.
 OS Porphyromonas gingivalis (Bacteroides gingivalis).
 OC Bacteria; CFB group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
 OC Porphyromonas.
 NCBI_TaxID=837;
 RN [1]
 RA SEQUENCE FROM N.A.
 RC STRAIN-W50;
 RX MEDLINE=99235907; PubMed=10219167;
 RA Stakeski N., Cleal S.M., Bhogal P.S., Reynolds E.C.;

RT *Characterization of a Porphyromonas gingivalis gene prtx that encodes
a lysine-specific cysteine proteinase and three sequence-related

RL Oral Microbiol. Immunol. 14:92-97(1999).

DR EMBL: U75366; AAB50809.1; -.

DR MEMOPS: C25.002; -.

DR InterPro: IPR000977; DNA_Ligase.

DR InterPro: IPR001769; Peptidase_C25.

DR Pfam: PF01364; Peptidase_C25; 3.

DR PROSITE: PS00697; DNA_LIGASE.A1; UNKNOWN.1.

SO SEQUENCE 1732 AA; 187914 MW; 45D5B91377391703 CRC64;

Query Match 100.0%; Score 78; DB 2; Length 1732;

Best Local Similarity 100.0%; Pred. No. 0.00013;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALNPDNYLSKDVYTG 15

DB 1216 ALNPDNYLSKDVYTG 1230

RESULT 13

ID 045985; PRELIMINARY; PRT; 334 AA.

DT 01-JUN-1998 (TREMBLrel. 06, Created)

DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE 2K1037.3 protein.

GN 2K1037.3.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;

OC Rhabditidae; Pelodierinae; Caenorhabditis.

OX NCBI_TaxID=6239;

RA none;

RP SEQUENCE FROM N.A.

RS Basam V.M.;

RL Submitted (Oct-1996) to the EMBL/GenBank/DBJ databases.

RN 12

RP SEQUENCE FROM N.A.

RX MEDLINE=99069613; PubMed=9851916;

RA none;

RT *Genome sequence of the nematode C. elegans: A platform for

investigating biology.";

RL Science 282:2012-2018(1998).

DR EMBL: Z81142; CA03502.1;

DR InterPro: IPR002851; DUF32.

DR Pfam: PF01748; DUF32; 1.

SO SEQUENCE 334 AA; 37719 MW; E8592D5A4D24FBA3 CRC64;

Query Match 60.3%; Score 47; DB 5; Length 334;

Best Local Similarity 69.2%; Pred. No. 5.9;

Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 LNPDNYLSKDVYTG 14

DB 22 LNPDNYLSEEDVT 34

RESULT 14

ID 09BH56

AC 09BH56; PRELIMINARY; PRT; 1579 AA.

DT 01-JUN-2001 (TREMBLrel. 17, Created)

DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)

DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)

DE M01G12.12 protein.

GN M01G12.12.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;

OC Rhabditidae; Pelodierinae; Caenorhabditis.

OX NCBI_TaxID=6239;

RA none;

RP SEQUENCE FROM N.A.

RP SEQUENCE FROM N.A.

RA McJury K.;

RL Submitted (Aug-1999) to the EMBL/GenBank/DBJ databases.

RN 12

RP SEQUENCE FROM N.A.

RX MEDLINE=99069613; PubMed=9851916;

RA none;

RT *Genome sequence of the nematode C. elegans: A platform for

investigating biology.";

RL Science 282:2012-2018(1998).

RN 13

RP SEQUENCE FROM N.A.

RA Lennard N.;

RL Submitted (Nov-1996) to the EMBL/GenBank/DBJ databases.

DR EMBL: AL132848; CAC35915.1; -.

DR EMBL: Z81571; CAC35915.1; JOINED.

DR EMBL: Z81571; CAC35914.1; -.

DR EMBL: AL132848; CAC35914.1; JOINED.

SO SEQUENCE 1579 AA; 182870 MW; B64517200715E26C CRC64;

Query Match 59.0%; Score 46; DB 5; Length 1579;

Best Local Similarity 46.7%; Pred. No. 47;

Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

OY 1 ALNPDNYLSKDVYTG 15

DB 601 SLGPDNYCLTHDLG 615

RESULT 15

ID 09KIB3; PRELIMINARY; PRT; 312 AA.

DT 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE Hypothetical outer membrane protein PG27.

OS Porphyromonas gingivalis (Bacteroides gingivalis).

OC Bacteria; CFB group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;

OC Porphyromonas.

OX NCBI_TaxID=837;

RA none;

RP SEQUENCE FROM N.A.

RS STRAIN-W50;

RL Ross B., Barr I., Patterson M., Agius C., Rothel L., Margelits M.,

Hocking D., Webb E.;

RT "P. gingivalis polypeptides and nucleic acids.";

RL Submitted (Feb-2000) to the EMBL/GenBank/DBJ databases.

RN 12

RP SEQUENCE FROM N.A.

RX STRAIN-W50;

RA Ross B.C.;

RT *Genomic analysis of Porphyromonas gingivalis for vaccine discovery.";

RL Submitted (Feb-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF237553; ABE81413.1; -.

DR InterPro: IPR001769; Peptidase_C25.

DR Pfam: PF01364; Peptidase_C25; 2.

SO SEQUENCE 312 AA; 34592 MW; 0D5792C9643A25P5 CRC64;

Query Match 57.7%; Score 45; DB 2; Length 312;

Best Local Similarity 61.5%; Pred. No. 12;

Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 ALNPDNYLSKDVYTG 13

DB 53 ALNPDNYLTPEL 65

Search completed: June 26, 2003, 00:53:25
Job time: 7.58824 secs

Thu Jun 26 11:58:08 2003

us-09-980-370-5.ra1

Page 2

```

FILING DATE: 25-JAN-1991
CLASSIFICATION: 424
PRIOR APPLICATION DATA: US 07/241,640
APPLICATION NUMBER: 08-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME (FILL IN): Ted M. 36, 965
RESIDENCE (FILL IN): UPLS-C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 372-5800
FAX: (904) 372-5800
INFORMATION FOR SEQ. ID NO.: 29
SEQUENCE CHARACTERISTICS:
LENGTH: 1687 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-570-311-29

Alignment Scores:
Pred. No.: 4,68e-77 Length: 1687
Score: 716.00 Matches: 131
Percent Similarity: 99.25% Conservative: 2
Best Local Similarity: 97.76% Mismatches: 1
Query Match: 97.95% Indels: 0
DB: 2 Gaps: 0

US-09-980-370-5 (1-402) x US-08-570-311-29 (1-1687)
OY 1 GCAGACTTCACGGAACGCTTCATCTGATGAGAGGACACGAGCAATGAGACT 60
DB 1120 AlasaphehrhrlurhpheluserSerthrhslgylalaproalaglurprrh 1139
OY 61 ACTATCGATGCCATGCGATGCGATGCGATGCGATGCGATGCGATGCGATGCGAT 120
DB 1140 ThrllaspalaaspolyaspolyglnlytrpuecySleuserSerlglylnleuasp 1159
OY 121 TGCTCACAGCTCATGCGCGACACGACGACGACGACGACGACGACGACGACGACG 180
DB 1160 Trpleuhtlrhlnslgyllylnhnsvalvalalaserphesertrpasmnglylmeala 1179
OY 181 TTGATCTGCTGATACATCTCATCTCAAGATCTTTCACGCGACGACGACGACGACG 240
DB 1180 Leuasnproaspasntylleuileserlyaspvalhtrhlyvalhlystlyr 1199
OY 241 TACTATCGATGACGACGCTTCCGCGGATGACGACGCTGATGACGACGACGACG 300
DB 1200 Tyrtlyralvalasmspelypheprogllyasphistylalvalhmetlleseulystr 1219
OY 301 GCGAGACGCGGACGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
DB 1220 Glythrasmnlaglyasphehrhvalvalphaglulurhproasnnglylleasnlys 1239
OY 361 GCGGAGACGATTCGCTTTCACGAGCAACGACGACGACGACGACGACGACGACG 402
DB 1240 GlyglyalalargpheclyleuserSerthrhslglnalasnnglyla 1253

RESULT 2
US-08-336-308A-10
Sequence 10, Application US/08336308A
GENERAL INFORMATION:
APPLICANT: Travis, James
APPLICANT: Potempa, Jan S.
APPLICANT: Barr, Philip J.
APPLICANT: Pavloff, Nadine
TITLE OF INVENTION: Porphyromonas gingivalis
TITLE OF INVENTION: Arginine-specific Proteinase Coding Sequences
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
Street: 5370 Manhattan Circle, Suite 201
```

```

CITY: Boulder
STATE: Colorado
COUNTRY: US
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/336,308A
FILING DATE: 08-NOV-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/119,361
FILING DATE: 10-SEP-1993
PRIOR APPLICATION DATA: US 08/265,441
APPLICATION NUMBER: 24-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Robert, Donna M. 33, 878
RESIDENCE (FILL IN): 33, 878
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 449-8080
FAX: (303) 449-8089
INFORMATION FOR SEQ. ID NO.: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1704 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-336-308A-10

Alignment Scores:
Pred. No.: 4,68e-77 Length: 1704
Score: 716.00 Matches: 131
Percent Similarity: 99.25% Conservative: 2
Best Local Similarity: 97.76% Mismatches: 1
Query Match: 97.95% Indels: 0
DB: 3 Gaps: 0

US-09-980-370-5 (1-402) x US-08-336-308A-10 (1-1704)
OY 1 GCAGACTTCACGGAACGCTTCGATGCTTCATCTGATGAGAGGACACGAGCAATGAGACT 60
DB 1137 AlasaphehrhrlurhpheluserSerthrhslgylalaproalaglurprrh 1156
OY 61 ACTATCGATGCCATGCGATGCGATGCGATGCGATGCGATGCGATGCGATGCGATG 120
DB 1157 ThrllaspalaaspolyaspolyglnlytrpuecySleuserSerlglylnleuasp 1176
OY 121 TGCTCACAGCTCATGCGCGACACGACGACGACGACGACGACGACGACGACGACG 180
DB 1177 Trpleuhtlrhlnslgyllylnhnsvalvalalaserphesertrpasmnglylmeala 1196
OY 181 TTGATCTGCTGATACATCTCATCTCAAGATCTTTCACGCGACGACGACGACGACG 240
DB 1197 Leuasnproaspasntylleuileserlyaspvalhtrhlyvalhlystlyr 1216
OY 241 TACTATCGATGACGACGCTTCCGCGGATGACGACGCTGATGACGACGACGACG 300
DB 1217 Tyrtlyralvalasmspelypheprogllyasphistylalvalhmetlleseulystr 1236
OY 301 GCGAGACGCGGACGACGCTTTCACGAGCAACGACGACGACGACGACGACGACGACG 360
DB 1237 Glythrasmnlaglyasphehrhvalvalphaglulurhproasnnglylleasnlys 1256
OY 361 GCGGAGACGATTCGCTTTCACGAGCAACGACGACGACGACGACGACGACGACG 402
DB 1257 GlyglyalalargpheclyleuserSerthrhslglnalasnnglyla 1270

RESULT 3
```

100

[REDACTED]

```

QY 61 ACTATGATGCCGATGGCATGTGAGGGTTGGCTGTCTGTCTTCCGCAATATGGAC 120
DB 1157 ThrileaspAlaaspGlyAspGlyGlnGlyTrpleucyLeuSerSerGlyGlnLeuasp 1176
QY 121 TGGCTCACAGCTCATGGCGGACCAAGCTAGTACTTCTTCATGGAATGGAATGGCT 180
DB 1177 TrpleuthrAlaHisGlyGlyThrAsnValValAlaIsePheSerTrpAsnGlyMetAla 1196
QY 181 TTGAATCCTGATTAATCTATCTCAAGGATGTTACAGCGCAACAAGTAAGTAC 240
DB 1197 LeuAsnProAspAsnTyrLeuIleSerLysAspValThrGlyAlaThrLysValLysTyr 1216
QY 241 TACTATCCAGTCAACGACGGTTTCCCGGATCACTATGCGGTATGATCTCCAAGC 300
DB 1217 TyrTyrAlaValAlaAsnAspGlyPheProGlyAspHisTyrAlaValMetIleSerLysThr 1236
QY 301 GGCAGAACGCCGAGACTTCAGCGTCTTTCAGAAAGCCCTAACGGAATTAATAG 360
DB 1237 GlyThrAsnAlaGlyAspPheThrValValAlaPheGlnGlnThrProAsnGlyIleAsnLys 1256
QY 361 GCGGAGCAAGATTGGCTCTTTCACAGCAAGCCAAATGGCGCC 402
DB 1257 GlyGlyAlaArgPheGlyLeuSerThrGlnAlaAsnGlyAla 1270

RESULT 5
US-08-570-311-10
; Sequence 10, Application US/08570311
; Patent No. 5824791
; GENERAL INFORMATION:
; APPLICANT: Proguiske-Fox, Ann
; APPLICANT: Tumwasorn, Somying
; APPLICANT: Lepline, Guyalaine
; APPLICANT: Han, Naibling
; APPLICANT: Lantze, Marillyn
; TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
; TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ted W. Whitlock
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/570,311
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/353,485
; FILING DATE: 09-DEC-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/647,119
; FILING DATE: 25-JAN-1991
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/241,640
; FILING DATE: 08-SEP-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Whitlock, Ted W
; REGISTRATION NUMBER: 36,965
; REFERENCE/DOCKET NUMBER: 0815.C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (904) 375-8100
; TELEFAX: (904) 372-5800
; INFORMATION FOR SEQ ID NO: 10:

```

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 1732 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-570-311-10

Alignment Scores:
Pred. No.: 6,2e-77 Length: 1732
Score: 715.00 Matches: 131
Percent Similarity: 99.254 Conservative: 2
Best Local Similarity: 97.764 Mismatches: 1
Query Match: 97.814 Indels: 0
DB: 2 Gaps: 0

US-09-980-370-5 (1-402) x US-08-570-311-10 (1-1732)
QY 1 GCAGACTTCACGGAACGCTTCAGCTCTTCTACTGATGAGAGCACCAGCGAATGCACT 60
DB 1157 ThrileaspAlaaspGlyAspGlyGlnGlyTrpleucyLeuSerSerGlyGlnLeuasp 1176
QY 61 ACTATGATGCCGATGGCATGTGAGGGTTGGCTGTCTGTCTTCCGCAATATGGAC 120
DB 1177 ThrileaspAlaaspGlyAspGlyGlnGlyTrpleucyLeuSerSerGlyGlnLeuasp 1196
QY 121 TGGCTCACAGCTCATGGCGGACCAAGCTAGTACTTCTTCATGGAATGGAATGACT 180
DB 1197 TrpleuthrAlaHisGlyGlyThrAsnValValAlaIsePheSerTrpAsnGlyMetAla 1216
QY 181 TTGAATCCTGATTAATCTATCTCAAGGATGTTACAGCGCAACAAGTAAGTAC 240
DB 1217 LeuAsnProAspAsnTyrLeuIleSerLysAspValThrGlyAlaThrLysValLysTyr 1236
QY 241 TACTATCCAGTCAACGACGGTTTCCCGGATCACTATGCGGTATGATCTCCAAGC 300
DB 1237 TyrTyrAlaValAlaAsnAspGlyPheProGlyAspHisTyrAlaValMetIleSerLysThr 1256
QY 301 GGCAGAACGCCGAGACTTCAGCGTCTTTCAGAAAGCCCTAACGGAATTAATAG 360
DB 1257 GlyThrAsnAlaGlyAspPheThrValValAlaPheGlnGlnThrProAsnGlyIleAsnLys 1276
QY 361 GCGGAGCAAGATTGGCTCTTTCACAGCAAGCCAAATGGCGCC 402
DB 1277 GlyGlyAlaArgPheGlyLeuSerThrGlnAlaAsnGlyAla 1290

RESULT 6
US-08-353-485-10
; Sequence 10, Application US/08353485
; Patent No. 5830710
; GENERAL INFORMATION:
; APPLICANT: Proguiske-Fox, Ann
; APPLICANT: Tumwasorn, Somying
; APPLICANT: Lepline, Guyalaine
; APPLICANT: Han, Naibling
; APPLICANT: Lantze, Marillyn
; TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
; TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ted W. Whitlock
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/353,485

```



```

: FILING DATE: 09-DEC-1994
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/647,119
: FILING DATE: 25-JAN-1991
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/241,640
: FILING DATE: 08-SEP-1988
: ATTORNEY/AGENT INFORMATION:
: NAME: Whitlock, Ted W.
: REGISTRATION NUMBER: 36,965
: REFERENCE/DOCKET NUMBER: UF15.C2
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (904) 375-8100
: TELEFAX: (904) 372-5800
: INFORMATION FOR SEQ ID NO: 10:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1732 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-353-485-10

Alignment Scores:
Pred. No.: 6,2e-77 Length: 1732
Score: 715.00 Matches: 131
Percent Similarity: 99.25% Conservative: 2
Best Local Similarity: 97.76% Mismatches: 1
Query Match: 97.81% Indels: 0
DB: 2 Gaps: 0

US-09-980-370-5 (1-402) x US-08-353-485-10 (1-1732)
QY 1 GCAGACTTCACGGAACGCTTCCTCTCATGAGAGCGACCGCATGAGCT 60
DB 1157 Alaasphethrcluhtrpnehlserthrhlsglylualaprohlaeglutrrpnr 1176
QY 61 ACTATGATGCCGATGCGGATGCGATGAGGCTTGCTGTCTCTCCGACAAATGGAC 120
DB 1177 Thrllasppalaasppilyaspdlglnclgtrpleucysleuserserglycineleuasp 1196
QY 121 TGGCTACAGACTCATGCGGCGCACCAAGCTAGTACTTTCATGAGTAATGAGGT 180
DB 1197 Trpleuthrhlahlselsglyserasnvalvalserpserhsertrrpsnnglymetala 1216
QY 181 TTGAATCCGTATACATCTCATCTCAAGATGTTACAGCGCCAGCAAGTAAGTAC 240
DB 1217 Leuasnprokspasmtyrleuileserlysaspvalthrlylatrhlrvsallystyr 1236
QY 241 TACTATCCAGTCACAGACGCTTTCCCGGATCATTATGCGGTGATCTCCAAAGACG 300
DB 1237 Tyrtfalaivalasnaasppilypneproclgysphlstryalvalmelleserlystr 1256
QY 301 GGCACGACGCGCGGACACTTCACGGTGTTCACAAAGACGCTTAAGCAATTAATAG 360
DB 1257 Glythrasmhlaeglyasphethrvalvalpnehlunlurhrproasnnglyleasnlys 1276
QY 361 GCGGACGCAAGATTCGGCTTTCCCGGACCAAGCCCAATGGCCCC 402
DB 1277 Glyglylalahrphnehlleuserthrlnlaasnnglyala 1290

RESULT 7
US-08-570-311-22
: Sequence 22, Application US/08570311
: Patent No. 5824791
: GENERAL INFORMATION:
: APPLICANT: Proguiske-Fox, Ann
: APPLICANT: Tummasorn, Somying
: APPLICANT: Lepline, Guyelaine
: APPLICANT: Han, Naiming
: APPLICANT: Lantz, Marilyn
: APPLICANT: Patti, Joseph

: TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
: AND PROBES FOR THE DETECTION OF PERIODONTAL DISEASE
: NUMBER OF SEQUENCES: 29
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Ted W. Whitlock
: STREET: 2421 N.W. 41st Street, Suite A-1
: CITY: Gainesville
: STATE: FL
: COUNTRY: USA
: ZIP: 32606
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/570,311
: FILING DATE:
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/353,485
: FILING DATE: 09-DEC-1994
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/647,119
: FILING DATE: 25-JAN-1991
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/241,640
: FILING DATE: 08-SEP-1988
: ATTORNEY/AGENT INFORMATION:
: NAME: Whitlock, Ted W.
: REGISTRATION NUMBER: 36,965
: REFERENCE/DOCKET NUMBER: UF15.C3
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (904) 375-8100
: TELEFAX: (904) 372-5800
: INFORMATION FOR SEQ ID NO: 22:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 439 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-570-311-22

Alignment Scores:
Pred. No.: 1.11e-76 Length: 439
Score: 711.00 Matches: 130
Percent Similarity: 99.25% Conservative: 3
Best Local Similarity: 97.01% Mismatches: 1
Query Match: 97.26% Indels: 0
DB: 2 Gaps: 0

US-09-980-370-5 (1-402) x US-08-570-311-22 (1-439)
QY 1 GCAGACTTCACGGAACGCTTCGCTTACTCATGAGAGCGACCGCATGAGCT 60
DB 192 Alaasphethrcluhtrpnehlserthrhlsglylualaprohlaeglutrrpnr 211
QY 61 ACTATGATGCCGATGCGGATGCGATGAGGCTTGCTGTCTCTCCGACAAATGGAC 120
DB 212 Thrllasppalaasppilyaspdlglnclgtrpleucysleuserserglycineleuasp 231
QY 121 TGGCTACAGACTCATGCGGCGCACCAAGCTAGTACTTTCATGAGTAATGAGGT 180
DB 232 Trpleuthrhlahlselsglythrasmhlaeglylualasrphesertrrpsnnglymetala 251
QY 181 TTGAATCCGTATACATCTCATCTCAAGATGTTACAGCGCCAGCAAGTAAGTAC 240
DB 252 Leuasnprokspasmtyrleuileserlysaspvalthrlylatrhlrvsallystyr 271
QY 241 TACTATCCAGTCACAGACGCTTTCCCGGATCATTATGCGGTGATCTCCAAAGACG 300
DB 271 Tyrtfalaivalasnaasppilypneproclgysphlstryalvalmelleserlystr 290
```

Page 6

Alignment Scores:	
Pred. No.:	1,12e-76
Score:	711.00
Percent Similarity:	99.25%
Best Local Similarity:	97.01%
Query Match:	97.26%
DB:	2
	Gaps:
	0

APPLICATION NUMBER: US 07/6647,113
FILING DATE: 25-JAN-1991
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/241,640
FILING DATE: 08-SEP-1988
ATTORNEY/AGENT INFORMATION:

NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UF15.C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 2628 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-570-311-14

Alignment Scores:
Pred. No.: 2,18e-76 Length: 2628
Score: 711.00 Matches: 130
Percent Similarity: 99.25% Conservative: 3
Best Local Similarity: 97.01% Mismatches: 1
Query Match: 97.26% Indels: 0
DB: 2 Gaps: 0

US-09-980-370-5 (1-402) x US-08-570-311-14 (1-2628)

QY 1 GCAGACTTCACGGAACGTTGAGTCTTCTACTCATGAGAGCAGCAGCGAATGAGCT 60
Db AlaaspphethrGluThrPhegluserSerThrHisGlyAlaIleProIaGluTrpThr 704
QY 61 ACATATGATGCCGATGCGGATGGTGAAGCTCTGCTGCTCTCTCCGACATTTGAC 120
Db ThrIleaspraIaaspGlyaspGlyGlnGlyTrpleucysleuserSerGlyGlnLeuasp 724
QY 121 TGCGTCACAGCTCATGCGCGCACCAACGTAGTAACTCTTTCATGATGAATGAGCT 180
Db TrrpleuThrAlaHisGlyGlyThrAsnValValAlaSerPheSerTrpAsnGlyMetAla 744
QY 181 TTGAATCTCGATTAACATCTCATCTCAAGAGATGTTACAGCGCACGAAGTAAGTAC 240
Db LeuasnProaspAsnTrpLeuIleSerLysaspValThrGlyAlaThrLysValLysTyr 764
QY 241 TACTATCCATCAACACGCTTTCCCGGGATCAGTACATGCGGTGATGATCTCCAACAG 300
Db TyTrpAlaValAlaAsnaspGlyAspHisTyrAlaValMetIleSerLysThr 784
QY 301 GGCAGCAACGCCGAGACTTCACGCTGTTTTCGAAGAAGCCCTAACGATAAATAAG 360
Db GlyThrAsnAlaGlyaspPheThrValAlaPhegluGluThrProAsnGlyIleAsnLys 804
QY 361 GCGGAGCAAGATTGCGTCTTTCCACGGAAGCCAAATGGCGCC 402
Db 805 GlyGlyAlaIaArgPheGlyLeuSerThrGluAlaIaaspGlyAla 818

RESULT 10
US-08-570-311-8
Sequence 8, Application US/08570311
Patent No. 5824791
GENERAL INFORMATION:
APPLICANT: Proquilske-Fox, Ann
APPLICANT: Tumwasorn, Somying
APPLICANT: Lepine, Guylaine
APPLICANT: Han, Naiming
APPLICANT: Lantz, Marilyn
APPLICANT: Patti, Joseph
TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ted W. Whitlock
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/570.311
FILING DATE:
CLASSIFICATION: 424
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/353,485
FILING DATE: 09-DEC-1994
CLASSIFICATION: 424
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/647,119
FILING DATE: 25-JAN-1991
CLASSIFICATION: 424
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/241,640
FILING DATE: 08-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UF15.C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1087 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-570-311-8

Alignment Scores:
Pred. No.: 2,71e-76 Length: 1087
Score: 709.00 Matches: 130
Percent Similarity: 98.51% Conservative: 2
Best Local Similarity: 97.01% Mismatches: 2
Query Match: 96.99% Indels: 0
DB: 2 Gaps: 0

US-09-980-370-5 (1-402) x US-08-570-311-8 (1-1087)

QY 1 GCAGACTTCACGGAACGTTGAGTCTTCTACTCATGAGAGCAGCAGCGAATGAGCT 60
Db AlaaspphethrGluThrPhegluserSerThrHisGlyAlaIleProIaGluTrpThr 539
QY 61 ACATATGATGCCGATGCGGATGGTGAAGCTCTGCTGCTCTCCGACAAATGAGCT 120
Db ThrIleaspraIaaspGlyaspGlyGlnGlyTrpleucysleuserSerGlyGlnLeuasp 559
QY 121 TGCGTCACAGCTCATGCGCGCACCAACGTAGTAACTCTTTCATGATGAATGAGCT 180
Db TrrpleuThrAlaHisGlyGlyThrAsnValValAlaSerPheSerTrpAsnGlyMetAla 579
QY 181 TTGAATCTCGATTAACATCTCATCTCAAGAGATGTTACAGCGCACGAAGTAAGTAC 240
Db LeuasnProaspAsnTrpLeuIleSerLysaspValThrGlyAlaThrLysValLysTyr 599
QY 241 TACTATCCATCAACACGCTTTCCCGGGATCAGTACATGCGGTGATGATCTCCAACAG 300
Db TyTrpAlaValAlaAsnaspGlyAspHisTyrAlaValMetIleSerLysThr 619
QY 301 GGCAGCAACGCCGAGACTTCACGCTGTTTTCGAAGAAGCCCTAACGATAAATAAG 360
Db GlyThrAsnAlaGlyaspPheThrValAlaPhegluGluThrProAsnGlyIleAsnLys 639
QY 361 GCGGAGCAAGATTGCGTCTTTCCACGGAAGCCAAATGGCGCC 402
Db 640 GlyGlyAlaIaArgPheGlyLeuSerThrGluAlaIaaspGlyAla 653

RESULT 11
US-08-353-485-8
Sequence 8, Application US/08353485
Patent No. 5830710
GENERAL INFORMATION:
APPLICANT: Proguiske-Fox, Ann
APPLICANT: Tummasorn, Somying
APPLICANT: Lepine, Guyalaine
APPLICANT: Han, Naiming
APPLICANT: Lantz, Marilyn
APPLICANT: Patti, Joseph
TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
TITLE OF INVENTION: and probes for the Detection of Periodontal Disease
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ted W. Whitlock
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/353,485
FILING DATE: 09-DEC-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/647,119
FILING DATE: 25-JAN-1991
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/241,640
FILING DATE: 08-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UF15.C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1087 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-353-485-8

Alignment Scores:
Pred. No.: 2,71e-76 Length: 1087
Score: 709.00 Matches: 130
Percent Similarity: 98.51% Conservative: 2
Best Local Similarity: 97.01% Mismatches: 2
Query Match: 96.99% Indels: 0
DB: 2 Gaps: 0

US-09-980-370-5 (1-402) x US-08-353-485-8 (1-1087)

OY 1 GCAGACTTCAGCGAAGCTTCGATCTCATGGAGGACGACGAGGAGTGGACT 60
DB 550 AAlAspRhetnRGltuThrPheGluSerSerThrnHsGluAlaProAlaGltuTrpThr 539
OY 61 ACTATCGATCGCGATGGGATGGGCTGCTGCTGCTCCGAGCAATTTGAGC 120
DB 540 ThrtLeAspAlaAspGlyAspGlyGlnAspTrpLeuGlyLeuSerSerGlyGlnLeuAsp 559
OY 121 TGGCTCAGACGCTCATGGCGGCGGACGACAGCTAGTAACTCTTTCTCATGGAATGGAGCT 180
DB 560 TrpleuThrAlaHnHsGlyGlyThrnAsnValValAlaLaserPheSerTrpAsnGlyMetAla 579

OY 181 TTGAATCCTGATTAATCATCTCATCAAGGATGTTACAGGCGCAAGGTAAGTAC 240
DB 580 LeuAsnProAspAsnTrpLeuIleSerLysAspValThrGlyAlaThrLysValLysTyr 559
OY 241 TACTATCAGCTCAGCAGCGTTTCCCGGGGATCAGTACGCGTATGATCTCCAGACG 300
DB 600 TyrTyrAlaValAsnAspGlyPheProGlyAspHisTyrAlaValMetLieserLysThr 619
OY 301 GGCAGCAGCGCGGAGCTTCACGTTGTTTTCGAGAAACGCTAAGGATAAATPAG 360
DB 620 GlyThrnAsnAlaGlyAspHetrnValValPheGluGluThrProAsnGlyIleAsnLys 639
OY 361 GGCAGCAGCAAGATTGCTCTTTCACGAGAACCCCAATGGCGCC 402
DB 640 GlyGlyAlaArpHeglyLeuSerTrnGluAlaAsnGlyAla 653

RESULT 12
US-08-570-311-27
Sequence 27, Application US/08570311
Patent No. 5824791
GENERAL INFORMATION:
APPLICANT: Proguiske-Fox, Ann
APPLICANT: Tummasorn, Somying
APPLICANT: Lepine, Guyalaine
APPLICANT: Han, Naiming
APPLICANT: Lantz, Marilyn
APPLICANT: Patti, Joseph
TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
TITLE OF INVENTION: and probes for the Detection of Periodontal Disease
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ted W. Whitlock
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/570,311
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/353,485
FILING DATE: 09-DEC-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/647,119
FILING DATE: 25-JAN-1991
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/241,640
FILING DATE: 08-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UF15.C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 1358 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-570-311-27

CORRESPONDENCE ADDRESS:

ADDRESSEE: Ted W. Whitlock
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/570,311
FILING DATE:

CLASSIFICATION: 424
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/353,485
FILING DATE: 09-DEC-1994

CLASSIFICATION: 424
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/647,119
FILING DATE: 25-JAN-1991

CLASSIFICATION: 424
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/241,640
FILING DATE: 08-SEP-1988

ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UF15.C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 372-5800

INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 456 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-570-311-20

Alignment Scores:

pred. No.: 7,74e-76 Length: 456
Score: 704.00 Matches: 129
Percent Similarity: 98.51% Conservative: 3
Best Local Similarity: 96.27% Mismatches: 2
Query Match: 96.31% Indels: 0
DB: 2 Gaps: 0

US-09-980-370-5 (1-402) x US-08-570-311-20 (1-456)

QY 1 GCAGACTCAGGAAACCTTGAGCTCTTCTCATGAGAGGACGACGAGGAATGACT 60
DB 192 AAlaAspPheHnGluThrPheGluSerSerThrHisGlyLualAProAlaGluTrpThr 211
QY 61 ACTATGATGCGGATGCGATGCTGAGGCTGCTGTCTGCTTCCGACAATTGAC 120
DB 212 ThrileAspAlaAspGlyAAspGlyGlnGlyTrpLeuGlySerSerGlyGlnLeuGly 231
QY 121 TGGCTCAGACGCTGAGGCGGACCACTACTAATCCTTCTCATGAGATGCAATGGCT 180
DB 232 TrpLeuTrpAlaHisGlyGlyThrAsnValAlaAlaSerPheSerTrpAsnGlyMetAla 251
QY 161 TTGAATCTGATTAACATCTCATCTCAAGAGATGTTACAGGCGACGAGAGTAAGTAC 240
DB 252 LeuAsnProAspAsnTyrLeuIleSerLysAspValThrGlyAlaThrLysValLysTyr 271
QY 241 TACTTCAGAGTCAGAGGAGTTTCCGGGATGAGTACATGAGTATGAGATCTCCAGACG 300
DB 272 TyrTrpAlaValAlaAsnAspGlyPheProGlyAspHisTyrAlaValMetIleSerLysThr 291
QY 301 GGCACCAAGCGCGAGACTTACGCGTTGTTTCAGAGAACAACGCTTAAGCAATTAATAG 360

DB 292 GlyThrAsnAlaGlyAAspPheThrValAlaPheGluGluTrpProAsnGlyIleAsnLys 311

QY 361 GCGGAGCAAGATTCGCTCTTCCAGGAGCAATGAGCGCC 402
DB 312 GlyGlyAlaArgPheGlyLeuSerThrGluAlaAspGlyAla 325

RESULT 15

US-08-570-311-2
Sequence 2, Application US/08570311
Patent No. 5824791

GENERAL INFORMATION:
APPLICANT: Proguiske-Fox, Ann
APPLICANT: Tumwasorn, Somying
APPLICANT: Lepine, Guylaine

APPLICANT: Han, Maiming
APPLICANT: Lantzi, Marilyn
APPLICANT: Patel, Joseph

TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:

ADDRESSEE: Ted W. Whitlock
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/570,311
FILING DATE:

CLASSIFICATION: 424
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/353,485
FILING DATE: 09-DEC-1994

CLASSIFICATION: 424
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/647,119
FILING DATE: 25-JAN-1991

CLASSIFICATION: 424
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/241,640
FILING DATE: 08-SEP-1988

ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UF15.C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 372-5800
TELEFAX: (904) 372-5800

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 497 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-570-311-2

US-09-980-370-5 (1-402) x US-08-570-311-2 (1-497)

Alignment Scores:

pred. No.: 9,55e-13 Length: 497
Score: 176.50 Matches: 45
Percent Similarity: 51.69% Conservative: 16
Best Local Similarity: 38.14% Mismatches: 44
Query Match: 24.15% Indels: 13
DB: 2 Gaps: 5

```

OY 13 GAAACGTTCCAGCTCTTCTACTCATGGAGAGCAGCCGGAATGCACTACTATCGATGCC 72
    |||:|||||:||||| ||| ||||| |||||
Db 235 GluSerPheAspThrGlnThr-----LeuProAsnGlyTrpThrMetIleAspAla 251
    |||:|||||:||||| ||| ||||| |||||
OY 73 GATGGCGATGATGAGGTTGGCTCTGCTCTTCCGACAATTGCACTGGCTCACAGCT 132
    |||:|||||:||||| ||| ||||| |||||
Db 252 AspGlyAspGlyHisAsnTrp-----LeuSerThrIleAsnValTyrAsnThrAlaThr 269
    |||:|||||:||||| ||| ||||| |||||
OY 133 CATGGCGGACCAACGAGTAAAGCTCTTCTCATGGATGGA-----ATG 177
    ||| ||| ||| ||| |||||
Db 270 HisThrGlyAspGlyAlaMetPheSerLysSerTrpThrAlaSerGlyGlyAlaLysIle 289
    ||| ||| ||| ||| |||||
OY 178 GCTTGAATCCTGATACTATCTCATCTCAAGATGTTACAGCGCCGCAAGAGGTAAG 237
    |||:|||||:||||| ||| ||||| |||||
Db 290 AspLeuSerProAspAsnTyrIleuValThrProLysValThrValProGluAsnGlyLys 309
    |||:|||||:||||| ||| ||||| |||||
OY 238 TACTACTATCCAGTCAACGACGAGGTTTCCC-----GGGATCACATATGCGGTGATGATC 291
    ||| ||| ||| ||| |||||
Db 310 LeuSerTyrTrpValSerSerGlnValProTrpThrAsnGluHisTyrGlyValPheLeu 329
    ||| ||| ||| ||| |||||
OY 292 TCCAGAGCGGCGCAGACGCGGAGACTTCACGGCT--GTTTCGAGAAGAAG 342
    ||| ||| ||| ||| |||||
Db 330 SerThrThrGlyAsnGluAlaAlaAsnPheThrIleLysLeuLeuGluGluThr 347
    ||| ||| ||| ||| |||||

```

Search completed: June 26, 2003, 00:47:33
 Job time : 36 secs

Thu Jun 26 11:58:09 2003

us-09-980-370-5.rapb

Page 2

```
Percent Similarity: 38.28% Conservative: 14
Best Local Similarity: 27.34% Mismatches: 48
Query Match: 9.75% Indels: 31
DB: 9 Gaps: 4
```

US-09-980-370-5 (1-402) X US-10-146-473-50 (1-2971),

Qy	40	GGCCGATGGCTTCGGTGAAAGACCAATCTGGTCGGCTTATTATTCGGTTGGC	342
Db	1162	SePProLeuAlaSerProValSerSerThrValSerValProLeuSerSerSerLeuPro	1181
Qy	341	GTCTTTCGAAAGAACCGGAAAGTCGCGGGCTTCGGCTTGGAATGATCAC	282
Db	1182	lIeSerValProThrThrLeuProAlaProAlaSerAlaProLeuThrIleProIleSer	1201
Qy	281	GCATPAGATGATCCCGGGAACCGTCTTACTGACTGATAGTAACTTAACTTCGTTGGC	222
Db	1202	Ala-----ProLeuThrValSerAlaSerGly-----	1210
Qy	221	CCTGTAACTCCTTTGGATGAGATAGTATGACGATTCAAAGCCATTCATTCATGAG	162
Db	1211	ProAlaLeu-----	1213
Qy	161	AAAGAGCTTACAGCTTGCGCGCATGAGCTGTGACGACCAATCTCGGAGAAC	102
Db	1214	-----LeuThrSerValThrProProLeuAlaProValValProAlaAlaProIlePro	1231
Qy	101	AGACAGAGCAACACCTCA--CCATCGCCATCGGCATGATGATGTCACCTTCCGCGTGGT	45
Db	1232	ProSerLeuAlaProSerGlyAlaSerProSerAlaSerAlaLeuThrLeuGlyLeuAla	1251
Qy	44	GCCTCTCATGATGATGAGACATCG	21
Db	1252	ThrAlaProSerLeuSerSerSer	1259

RESULT 2
US-09-738-626-4263
Sequence 4263, Application US/09738626

PUBLIC ID NO: US20020197003A
 GENERAL INFORMATION:
 APPLICANT: NAKAGAMI, SATOSHI
 APPLICANT: MIYAZUKI, HIROSHI
 APPLICANT: NAGAI, SHIRO
 APPLICANT: NAGAI, SHIRO
 APPLICANT: OHNAMI, KENICHI
 APPLICANT: YOKOI, HARUHIKO
 APPLICANT: TATEISHI, MAMOKO
 APPLICANT: SENOH, AKIHITO
 APPLICANT: IKEDA, MASATO
 APPLICANT: OZAKI, AKIO
 TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 FILE REFERENCE: 249-125
 CURRENT APPLICATION NUMBER: US/09/738, 626
 CURRENT FILING DATE: 2000-12-18
 PRIOR APPLICATION NUMBER: JP 99/377484
 PRIOR FILING DATE: 1999-12-06
 PRIOR APPLICATION NUMBER: JP 00/159162
 PRIOR FILING DATE: 2000-04-07
 PRIOR APPLICATION NUMBER: JP 00/280988
 PRIOR FILING DATE: 2000-08-03
 NUMBER OF SEQ ID NOS: 7059
 SOFTWARE: patentin ver. 3.0
 SEQ ID NO 4263

: ORGANISM:	Corynebacterium glutamicum
US-09-738-626-4263	
 Alignment Scores:	
Pred. No.:	7..08
Score:	73..00
Percent Similarity:	40.85%
Best Local Similarity:	28.66%

Query Match:	9.99%	Indels:	40
DB:	9	Gaps:	9

US-09-980-370-5 (1-402) X US-09-738-626-4263 (1-575)

OY	8	TCACGAAGACGTTCCAGCTCTTTACTCATGGAGAGCACCACGGGATAGCATATACG	67
Db	27	SeRALaValaIoIuSerSerIuElIeValIleRoIaRgIySgImaRgIaIdElYalPheMet	46
OY	68	ATGCCGATGGCGATGCGTAGCGTTGGCTGTGTCTTCGCGACA---ATGGACTGCG	124
Db	47	ASpGIuSerAraGloIuSerPheGlyThrIaGlyIeu---ArgAlaProValGlyProI	66
OY	125	TCACAGCTCATGGCGCACCAACGTAAGTACGTC	158
Db	66	a-----ArgHisGImeCAsValIeuGlnValIleAraGlyThrIaGlyIa	82
OY	159	-----TTTCTCATGAGATAGCATGGCTTTGATCTGATATATCTCATAGG	211
Db	82	IaIaSerPheIuElIaGlyI---ArgAlaIaIeIuSuInProValProHisIeuValIProGIIa	102
OY	212	ATGTTCACAGCGCCACGAAGGTAAAGTACTACATTCATCCAGTCAC	255
Db	102	SpIuThrGlyIleGlyIeGlyAraIa-----LeuIyIProGlnaIaSpGlyProIeUaRgValIy	120
OY	256	-----GACGGTTTTCCGGGGGATCAC---TATCGCGTATGATCTCCAAACGG	301
Db	120	aIValaGlyTyrAspAlaIaRgIyGlyGlySerHisThrPheAlaIaIaThrIleThrIaGlyIaIeP	140
OY	302	GCACGAACGCCGAGACCTTCACGGTTGTTCGAAAGAACGCCCTAAACGSA-	351
Db	140	healIaGlyIaGlyI---PheIuIValIleThrIeuleuPProIleProSerProIleProluI	159
OY	352	-----ATAATTAAGSGGCGACGACGAAGATTCGGCTTTCCACGGAGACC-----A	394
Db	159	IeProTPIeuValaIaSuIlyHisGlyIeuaSpAlaIdElYalGlnIleThrIaIeSerHisA	179
OY	395	ATGGCGGCC	402
Db	179	SmIyAla	181

RESULT 3
US-09-866-050A-303
; Sequence 303, Application US/09866050A

```

1 PDBID:INFO:09-866-050A/1AI
2 GENERAL INFORMATION:
3 APPLICANT: Watson, James D.
4 APPLICANT: Strachan, Lorna
5 APPLICANT: Sleeman, Matthew
6 APPLICANT: Onrust, Renee
7 APPLICANT: Marison, James G.
8 APPLICANT: Kumble, Krishnand D.
9 TITLE OF INVENTION: Compositions Isolated from Skin Cells
10 FILE REFERENCE: 11000.1011c4U
11 CURRENT APPLICATION NUMBER: US/09/866.050A
12 CURRENT FILING DATE: 2001-05-24
13 NUMBER OF SEQ ID NOS: 725
14 SOFTWARE: FASTSEQ for Windows Version 4.0
15 SEQ ID NO 303
16 LENGTH: 617
17 TYPE: PRT
18 ORGANISM: Mouse
19 US-09-866-050A-303

```

Alignment Scores:	
Pred. No.:	8.29
Score:	617
Matches:	72.50
Percent Similarity:	41.33%
Best Local Similarity:	29.33%
Query Match:	9.92%
DB:	9
Length:	617
Conservative:	22
Mismatches:	13
Indels:	3
Caps:	3

US-09-980-370-5 (1-402) x US-09-866-050A-303 (1-617)

QY 7 TTCACGGAACGTTGAGTCTTACTACTGAGAGGACGACGAGTGGACTACTATC 66
DB 301 PhelypheserPhelulaserPhelulaserPhelulaserPhelulaser 216
QY 67 GATCCGATGCGATGCGATGCGATGCGATGCGATGCGATGCGATGCGATGCGATC 126
DB 317 -----SerGlyserLeuProSerTrpThrValSerMetAlaPheGlnGlnGln 334
QY 127 ACAGCTATGCGGCGACGACGATGAGTCTTCTCA-----TGCATGGAATGAGCT 180
DB 335 ThreLeuHisPheSerSerTrpThrValSerMetAlaPheGlnGlnGln 354
QY 181 TTGATCTGATTAAC-----TATCTCATCTCAAG 210
DB 355 LeuGlyProAspThrProMetProProValTyrSerLeuSerGln 369

RESULT 4
US-08-424-550B-82
Sequence 82 Application US/08424550B
Patent No. US2002019447A1
GENERAL INFORMATION:
APPLICANT: JOHN N. SIMONS
APPLICANT: TAMM J. PILOT-MATIAS
APPLICANT: GEORGE J. DAMSON
APPLICANT: GEORGE J. SCHLAUDER
APPLICANT: GURUSH R. DESAI
APPLICANT: JAMES C. ERKER
APPLICANT: ANTHONY SCOTT MUEHROF
APPLICANT: SHERI L. BUIK
APPLICANT: ISA K. MUSHAMMAR
TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE
NUMBER OF SEQUENCES: 16
SEQUENCE LENGTH: 16
ADDRESSER: ABBOTT LABORATORIES D377/A86D
STREET: 100 ABBOTT PARK ROAD
CITY: ABBOTT PARK
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/424,550B
FILING DATE: 08/93/8-2633
CLASSIFICATION: 435435
ATTORNEY/AGENT INFORMATION:
NAME: POREMBSKI, PRISCILLA E.
REGISTRATION NUMBER: 33,207
REFERENCE/DOCKET NUMBER: 5527.PC.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 708-937-6365
TELEFAX: 708-938-2633
INFORMATION FOR SEQUENCING: 82:
SEQUENCE CHARACTERISTICS:
LENGTH: 1422 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-424-550B-82

Alignment Scores:
Pred. No.: 13.9 Length: 1422
Score: 71.50 Matches: 29
Percent Similarity: 34.91% Conservative: 8
Best Local Similarity: 27.36% Mismatches: 34

Query Match: 9.78% Indels: 35
DB: 8 Gaps: 7

US-09-980-370-5 (1-402) x US-08-424-550B-82 (1-1422)

QY 30 TATCATGAGAGAGGACGACGACGACGACGACGACGACGACGACGACGACGACG 89
DB 556 TYPHETIPASPSERATGTHR-----ArgCys---TrpArgTrpSerGly 569
QY 90 TTGCGTCTGCTGCTTCCGACAAATGAGTGGCTGACACTCA---TGGCGGACGAA 146
DB 570 TYRGLYLeuSer-----SerHis**HisPheTrpArgHisLeu 582
QY 147 CTTAGTAAGCTTTCTTCATGATGATGATGATGATGATGATGATGATGATGATGATC 206
DB 583 CysAlaAlaLeuLeuValTyrTyrLeuSerProTyrArg----- 595
QY 207 AAAGATGTTTACAGCGCGACGACGACGACGACGACGACGACGACGACGACGACG 266
DB 596 -----CysTyrCysArgProSerGly-----ArgArgGlyAsn 607
QY 267 CGGGATCACTATGCGGTATGATCTCAAGACGACGACGACGACGACGACGACGACG 326
DB 608 ArgGlyGlyValCysIleIleHisSerLeuGlyGly-----HisGly 621
QY 327 TGTCTTTCGAAAGAACGCC 344
DB 622 CysGlyAsn**GlnAla 627

RESULT 5
US-09-784-877-2
Sequence 2 Application US/09784877
Patent No. US20010012836A1
GENERAL INFORMATION:
APPLICANT: Exding Hu
APPLICANT: Yuan Zhu
APPLICANT: YUAN ZHU
TITLE OF INVENTION: HUMAN HISTONE DEACETYLASE GENE HD4
TITLE OF INVENTION: GP-70516-C1
CURRENT APPLICATION NUMBER: US/09/784,877
PRIOR FILING DATE: 2001-02-16
CURRENT FILING DATE: 1998-08-21
NUMBER OF SEQ ID NOS: 2
SEQUENCE LENGTH: 377
SEQUENCE ID NO: 2
SEQUENCE FASTSEQ FOR Windows Version 3.0
LENGTH: 377
TYPE: PRT
ORGANISM: HOMO SAPIENS
US-09-784-877-2

Alignment Scores:
Pred. No.: 14.3 Length: 377
Score: 70.00 Matches: 32
Percent Similarity: 23.09% Conservative: 13
Best Local Similarity: 23.53% Mismatches: 45
Query Match: 9.58% Indels: 46
DB: 10 Gaps: 5

US-09-980-370-5 (1-402) x US-09-784-877-2 (1-377)

QY 124 CTACACGCTATGCGGCGACGACGACGACGACGACGACGACGACGACGACGACG 183
DB 177 LeuAspLeuHisIleGlyAspValIleGluAspAlaPheSerPheThrValMet 196
QY 184 -----TATCTGATTAATCTATCTCAAGAGATGTTTACA--- 219
DB 197 ThreValSerLeuHisLysPheSerTrpGlyPheThrGlyThrValSerValSerAsp 216
QY 220 -----GGCGGACGACGACGACGACGACGACGACGACGACGACGACGACG 267
DB 217 ValGlyLeuGlyLysGlyArgTyrTyrSerValAsnValProIleGlnAspGlyIleGln 236

```
OY 268 GGGATCAGTCTAT----- 279
Db 237 AspGluuTyTyrGlnIleCysGluSerValLeuYsgluValTyGlnAlaPheasn 256
OY 280 -----GGCGTGATGATCTCCAGACGGCGACCGCCGAGACTTCACGGTTGTT 330
Db 257 ProluAlaValValLeuGlnLeuGlnAlaAspThrIleAlaGlyAspPrometCysSer 276
OY 331 TTGCAAGAAAGCCCTAACGGAATAAATAG----- 360
Db 277 PheasnMetThrProValGlyIleGlyCysLeuYsTyTyleuGlnTrpGlnLeu 296
OY 361 -----GGCGAGCAAGATTCGGCTTTCCAGCGAAGCC 393
Db 297 AlaThrLeuIleLeuGlyGlyGlyGlyTyrAsnLeuAlaAsnThrAla 312

RESULT 6
US-09-817-913-15
: Sequence 15, Application US/09817913
: Patent No. US20020061860A1
: GENERAL INFORMATION:
: APPLICANT: L.L. Zucmel, Claire
: APPLICANT: Bonfils, Jeffrey
: APPLICANT: Besterman, Jeffrey
: TITLE OF INVENTION: Inhibition of Specific Histone Deacetylase Isoforms
: FILE REFERENCE: 106101.145
: CURRENT FILING DATE: 2001-03-26
: PRIOR APPLICATION NUMBER: US 60/192,157
: NUMBER OF SEQ ID NOS: 33
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 15
: LENGTH: 377
: TYPE: PRN
: ORGANISM: Human
US-09-817-913-15

Alignment Scores:
Pred. No.: 14.3 Length: 377
Score: 70.00 Matches: 32
Percent Similarity: 33.09% Conservative: 13
Best Local Similarity: 23.58% Mismatches: 45
Query Match: 9.58% Indels: 46
Gaps: 5
DB: 10

US-09-980-370-5 (1-402) x US-09-817-913-15 (1-377)
OY 124 CTCACAGCTCATGCGGACCAAGCTAGTCTTTCTCATGAGATGAGTGGCTTG 183
Db 177 LeuAspLeuIleHisIleAspGlyValGluAspAlaPheSerThrSerValMet 196
OY 184 -----AATCTGTATACCTATCTCATGCTCAAGAGATGTTACA--- 219
Db 197 ThrValSerLeuHisIleYsPheSerProGlyPhePheProGlyThrGlyAspValSerAsp 216
OY 220 -----GGCGCAAGCAAGTAAAGTACTAC-----TATCCAGTCAAGCAAGCTTTTCCC 267
Db 217 ValGlyLeuGlyLeuGlyArgTyTyrSerValAsnValProIleGlnAspGlyIleGln 236
OY 268 GGGATCAGTCTAT----- 279
Db 237 AspGluuTyTyrGlnIleCysGluSerValLeuYsgluValTyGlnAlaPheasn 256
OY 280 -----GGCGTGATGATCTCCAGACGGCGACCGCCGAGACTTCACGGTTGTT 330
Db 257 ProluAlaValValLeuGlnLeuGlnAlaAspThrIleAlaGlyAspPrometCysSer 276
OY 331 TTGCAAGAAAGCCCTAACGGAATAAATAG----- 360
Db 277 PheasnMetThrProValGlyIleGlyCysLeuYsTyTyleuGlnTrpGlnLeu 296
OY 361 -----GGCGAGCAAGATTCGGCTTTCCAGCGAAGCC 393
```

```
Db 297 AlaThrLeuIleLeuGlyGlyGlyGlyTyrAsnLeuAlaAsnThrAla 312

RESULT 7
US-09-817-538-15
: Sequence 15, Application US/09817538
: Patent No. US20020137162A1
: GENERAL INFORMATION:
: APPLICANT: Bonfils, Claire
: APPLICANT: Besterman, Jeffrey
: APPLICANT: Besterman, Jeffrey
: TITLE OF INVENTION: Antisense Oligonucleotide Inhibition of Specific Histone
: FILE REFERENCE: 106101.144
: CURRENT FILING DATE: 2001-03-26
: PRIOR APPLICATION NUMBER: US 60/192,157
: NUMBER OF SEQ ID NOS: 33
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 15
: LENGTH: 377
: TYPE: PRN
: ORGANISM: Human
US-09-817-538-15

Alignment Scores:
Pred. No.: 14.3 Length: 377
Score: 70.00 Matches: 32
Percent Similarity: 33.09% Conservative: 13
Best Local Similarity: 23.58% Mismatches: 45
Query Match: 9.58% Indels: 46
Gaps: 5
DB: 10

US-09-980-370-5 (1-402) x US-09-817-538-15 (1-377)
OY 124 CTCACAGCTCATGCGGACCAAGCTAGTCTTTCTCATGAGATGAGTGGCTTG 183
Db 177 LeuAspLeuIleHisIleAspGlyValGluAspAlaPheSerThrSerValMet 196
OY 184 -----AATCTGTATACCTATCTCATGCTCAAGAGATGTTACA--- 219
Db 197 ThrValSerLeuHisIleYsPheSerProGlyPhePheProGlyThrGlyAspValSerAsp 216
OY 220 -----GGCGCAAGCAAGTAAAGTACTAC-----TATCCAGTCAAGCAAGCTTTTCCC 267
Db 217 ValGlyLeuGlyLeuGlyArgTyTyrSerValAsnValProIleGlnAspGlyIleGln 236
OY 268 GGGATCAGTCTAT----- 279
Db 237 AspGluuTyTyrGlnIleCysGluSerValLeuYsgluValTyGlnAlaPheasn 256
OY 280 -----GGCGTGATGATCTCCAGACGGCGACCGCCGAGACTTCACGGTTGTT 330
Db 257 ProluAlaValValLeuGlnLeuGlnAlaAspThrIleAlaGlyAspPrometCysSer 276
OY 331 TTGCAAGAAAGCCCTAACGGAATAAATAG----- 360
Db 277 PheasnMetThrProValGlyIleGlyCysLeuYsTyTyleuGlnTrpGlnLeu 296
OY 361 -----GGCGAGCAAGATTCGGCTTTCCAGCGAAGCC 393
Db 297 AlaThrLeuIleLeuGlyGlyGlyGlyTyrAsnLeuAlaAsnThrAla 312

RESULT 8
US-09-308-207-33
: Sequence 33, Application US/09308207
: Patent No. US2003002233A1
: GENERAL INFORMATION:
: APPLICANT: MARIA DIAZ-TORRES ET AL.
: TITLE OF INVENTION: METHOD FOR THE RECOMBINANT
: PRODUCTION OF 1,3 PROPANEDIOL
: NUMBER OF SEQUENCES: 68
```

CORRESPONDENCE ADDRESS:
 ADDRESSEE: Genencor International, Inc.
 STREET: 4 Cambridge Place
 CITY: Rochester
 STATE: NY
 COUNTRY: U.S.A
 ZIP: 14618
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: Windows
 SOFTWARE: FastSeq for Windows Version 2.0b
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/308,207
 FILING DATE: 13-May-1999
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/030,601
 FILING DATE: 13-NOV-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Glaisier, Debra
 REGISTRATION NUMBER: 33,888
 REFERENCE/DOCKET NUMBER: GC 369-2
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 650-864-7620
 TELEFAX: 650-845-6504
 TELEX: <Unknown>
 INFORMATION FOR SEQ ID NO: 33:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 271 amino acids
 TYPE: amino acid
 STRANDEDNESS: unknown
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 ORIGINAL SOURCE:
 ORGANISM: Gppl
 SEQUENCE DESCRIPTION: SEQ ID NO: 33:
 US-09-308-207-33
 Alignment Scores:
 Pred. No.: 17.1 Length: 271
 Score: 69.00 Matches: 33
 Percent Similarity: 40.44% Conservative: 22
 Best Local Similarity: 24.26% Mismatches: 53
 Query Match: 9.44% Indels: 28
 Gaps: 7
 US-09-980-370-5 (1-402) x US-09-308-207-33 (1-271)
 QY 4 GACTTCAGC--GAAACGTTGAGTCTTCTCATGAGAGCGACACCGAGCAATGGACT 60
 DB 89 AASPHEALAAAPGLUGLUTYRVALAENLYSLEUGLUGLUTLEPROGLWUYSTYRGly 108
 QY 61 ACTATGATGCCGATGGCATGTGAGGGTTGGCTGTCTCTCTCCGACAAATGG--- 117
 DB 109 GIUHLISERTILEGLVALAIPROGLYALVALLYSLEUCYSANALALEUANSALALEUPRO 128
 QY 118 -----GACTGGCTCAGCATGCGCGCACCAACGATGATGAGCTTTCTTCATGAGAT 171
 DB 129 LYSGLULSTPRALVALAALHRSERGLYHRAASPMETALALYSLSYSTRPHEASP 148
 QY 172 GGAATGCGCTTGG---AATCTGATATACTATCTCATCTCAAGAGATGT---ACAGCGCA 225
 DB 149 ILELEULIILELALYARPROGLUTYRPHETLEHRAALASNPVALLYSGLINGLYLS 168
 QY 226 ACGAAGTAAGTACTAC-----TATCCAGTCAACGACGGT 261
 DB 169 PROHLSPROGLIPROTYLLEULYSELARASNGLYLEUGLYPHEPROILEASNGLYLN 188
 QY 262 TTTCCCGGGGATCACTATGCGGTGATGATCTCAAGACGGCACGAACCGCGAGACTTC 321
 DB 189 ASPRO-----SerLYSserLYS----- 194

QY 322 ACGTTGTTTCGAAGAACCCCTACGGAATTAATAGCGCGAGCA 369
 DB 195 VALVALVALPHEGLVASPALAPROALGLYLEALALAGLYLYSALA 210
 RESULT 9
 US-09-924-256A-46
 : Sequence 46, Application US/09924256A
 : Patent No. US20020127659A1
 : GENERAL INFORMATION:
 : APPLICANT: Waters, Barbara
 : APPLICANT: Miao, Vivian
 : APPLICANT: Ho, Yap
 : APPLICANT: Tong, Seow
 : TITLE OF INVENTION: METHOD FOR ISOLATION OF BIOSYNTHESIS GENES FOR
 : FILE REFERENCE: 9993-006
 : CURRENT APPLICATION NUMBER: US/09/924,256A
 : PRIOR FILING DATE: 2001-08-08
 : PRIOR APPLICATION NUMBER: 08/861,774
 : NUMBER OF SEQ ID NOS: 94
 : SOFTWARE: PatentIn Ver. 2.0
 : SEQ ID NO 46
 : LENGTH: 218
 : TYPE: PRT
 : ORGANISM: Parmelia sulcata
 : FEATURE:
 : NAME/KEY: SITE
 : LOCATION: (94)
 : OTHER INFORMATION: Xaa=unknown amino acid
 US-09-924-256A-46
 Alignment Scores:
 Pred. No.: 21.2 Length: 218
 Score: 68.00 Matches: 39
 Percent Similarity: 39.31% Conservative: 18
 Best Local Similarity: 26.90% Mismatches: 51
 Query Match: 9.30% Indels: 37
 DB: 10 Gaps: 8
 US-09-980-370-5 (1-402) x US-09-924-256A-46 (1-218)
 QY 6 CTCACGGAAGCTGAGTCTTCTCATGAGAGCGACACCGCAATGGACTACTAT 65
 DB 60 LEUSERSANRGLIESERTYRPHENPEASNTRPHISGLYALASERILEHRLMETASPTHR 79
 QY 66 CGATGCCGATGGCGATGAGTGGTGGCTGTCTGCTCTCCGACAAATTGGACTGGCT 125
 DB 80 ALACYSER-----SerSERLEUVALALALILEHLSLEUVALA 91
 QY 126 CACAGCTCATGCGCGCACCAACGATGATGACTCTTTCTCATGGAATGA-----ATGCC 179
 DB 92 VALGLN**LEUAFRGLAASNGIUSERAFRGMEVALAVALA-CYSGLYSERASNLEUI 111
 QY 180 TTGAATCCGATAACTATCTCATCAAGATGTTACAGGGCGCAACGATTAAGTA 239
 DB 111 ELEUGLYPROGLINSERTYRILELEGLU-----SerLYSVALLYSME 125
 QY 240 CTACTATCCAGTCACGACGCGTTTCC----- 267
 DB 125 LLEUSERPRO-----ASPLYLEUSERARGMETTRPAPLYASPALASNGLYTYRAL 143
 QY 268 -----GGGATCATATGCGGTGATGATCTTCACAGAGCGGACGACAGCC-----GG 314
 DB 143 AARGGLYASPGLYVALALAVALLAVALLEULYSTRHLEUSERVALALALEUVALASPGI 163
 QY 315 AGAC--TTTACCGGTGTTTTCGAAGAAGCGCTTAACGGAATAATAAGCGCGAGCAG 371
 DB 163 YASPHSLILEGLUCYCLEULEARGLTHR-----GLYLEUASNGINASPLYALATH 181
 QY 372 ATTGCGTCTTTC 384


```

OY 136 GCGCGCACCAACAGTAGTAGCTTTTCATGAGTAGTAGGCTTTGATCTGATTAAC 195
    ||| ::::
DB 636 GlysanserGlnite----- 640
OY 196 TATCTCATCTCAAGAGATTGTACAGCGGCAACAGTAGTAGTACTGATCTGATTAAC 255
    ||| ::::
DB 641 -----ValArgAlaTyrValProLeuAla 648
OY 256 GACGGTTTTCCCGGAGTACTATGCGGTGATGATCTCAAGACGCGGACGACGCGCGA 315
    :::: ||| ::::
DB 649 GlnMetPhe-----GlyTyrAlaThrValLeuArgSerAlaSerGlnGlyArgGly 665
OY 316 GACTTCACAGCTTGTTTC-----GAGCAACGCGCTTAAGCAATTA 354
    ||| ::::
DB 666 ThrPheMetMetValPheAspHisTyrGlnAspValProLysSerValGlnGlnGln 685
OY 355 -----AATTAAGGC 363
    ||| ::::
DB 686 IleLysLysAsnLysGly 691

RESULT 12
US-10-125-692-18
: Sequence 18, Application US/10125692
: Publication No. US20030044429A1
: GENERAL INFORMATION:
: APPLICANT: Adem, Alan
: APPLICANT: Hayashi, Fumitaka
: APPLICANT: Smith, Kelly D.
: APPLICANT: Underhill, David M.
: APPLICANT: Ozlinsky, Adrian
: TITLE OF INVENTION: Toll-Like Receptor 5 Ligands and Methods
: FILE REFERENCE: P-15, 5155
: CURRENT APPLICATION NUMBER: US/10/125,692
: PRIOR FILING DATE: 2002-04-17
: PRIOR APPLICATION NUMBER: US 60/285,477
: PRIOR FILING DATE: 2001-04-20
: NUMBER OF SEQ ID NOS: 43
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 18
: LENGTH: 506
: TYPE: PR
: ORGANISM: S. typhimurium2
US-10-125-692-18

Alignment Scores:
Pred. No.: 31 Length: 506
Score: 67.50 Matches: 36
Percent Similarity: 37.88% Conservative: 14
Best Local Similarity: 27.27% Mismatches: 53
Query Match: 9.23% Indels: 29
DB: 9 Gaps: 6

US-09-980-370-5 (1-402) x US-10-125-692-18 (1-506)
OY 58 ACTACTATGATGCCGATGCGGATGCGGCTGCTGCTGCTCGGACAAATG 117
    ||| ::::
DB 196 ThrThrLeuAspValSerGlyLeuAspAlaAla----- 208
OY 118 GACTGGCTCACAGCTCATGCGGCGGACCAACAGTAGTAGCTTTTCATGATGAATG 177
    ||| ||| ||| |||
DB 209 -----LysAlaAlaThrGlyGlyThrAsnGlyThrAlaSerValThrGlnGlyAlaVal 226
OY 178 GCTTGATCTCGATTAAC-----TATCTCATCTCAAGAGATTGTACAGCGCGA 228
    :::: ||| ::::
DB 227 LysPheAspAlaAspAsnAsnLysTyrPheValThrIleGlyGlyPheThrGlyAlaAsp 246
OY 229 AAGTAAGTAC---TACTATCCAGTCAACAGCGGCTTTCCCGGAGT---CACTATGC 282
    ||| ||| ||| |||
DB 247 AlaAlaLysAsnGlyAspTyrGlnValAsnValAlaThrAspGlyThrValThrLeuAla 266
OY 283 GTCAGATCTCCAAAGACGGGACGAAACCGCGGAGACTTCACG----- 324

```

```

DB 267 AlaGlyAlaThrLysThrThrMetProAlaGlyAlaThrThrLysThrGlnValGlnGlu 286
    :::: ||| |||
OY 325 -----GTTGTTTGAAGAAAGCCTTACGAGTAATAAAGCGCGA 366
DB 287 LeuLysAspThrProAlaValAlaValSerAlaAspAlaLysAsnAlaLeuIleAlaGlyGly 306
OY 367 GCAGATTGCGTCTTCACGGAAGCCCAATGCGGC 402
    :::: ||| ::::
DB 307 ValAsp-----AlaThrAspAlaAsnGlyAla 315

RESULT 13
US-09-980-748-1745
: Sequence 1745, Application US/09980748
: Publication No. US20030059937A1
: GENERAL INFORMATION:
: APPLICANT: Ruben et al.
: TITLE OF INVENTION: Antipodites that Immunospecifically Bind Blys
: FILE REFERENCE: P6523
: CURRENT APPLICATION NUMBER: US/09/880,748
: PRIOR FILING DATE: 2001-06-15
: PRIOR APPLICATION NUMBER: 60/212,210
: PRIOR FILING DATE: 2000-06-15
: PRIOR APPLICATION NUMBER: 60/240,816
: PRIOR FILING DATE: 2000-10-17,248
: PRIOR APPLICATION NUMBER: 60/276,248
: PRIOR FILING DATE: 2001-03-16
: PRIOR APPLICATION NUMBER: 60/277,379
: PRIOR FILING DATE: 2001-03-21
: PRIOR APPLICATION NUMBER: 60/293,499
: PRIOR FILING DATE: 2001-05-25
: NUMBER OF SEQ ID NOS: 3239
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 1745
: LENGTH: 256
: TYPE: PR
: ORGANISM: Homo sapiens
US-09-980-748-1745

Alignment Scores:
Pred. No.: 29.2 Length: 256
Score: 67.00 Matches: 31
Percent Similarity: 33.12% Conservative: 21
Best Local Similarity: 19.75% Mismatches: 49
Query Match: 9.17% Indels: 56
DB: 9 Gaps: 7

US-09-980-370-5 (1-402) x US-09-980-748-1745 (1-256)
OY 46 CCAGCGAATGAGTACTATGATGCGGATGCGGAGTGGCTGCTGCTGCT 105
    ||| ::::
DB 14 ProSerGlnThrLeuSerLeuThrCysThrValSerGlySer-----IleSer 30
OY 106 TCCGACAAATG-----GACTGGCTCACAGCTCATGCGGCGGACCAACGTAAGTCTCT 159
    ||| ||| ||| |||
DB 31 SerGlyAsnTyrTyrTTPSerTTPValArgGlnHisProIly-----LysGly 46
OY 160 TTCTCATGAATGCA----- 174
    ||| |||
DB 47 LeuGlnTTPIleGlyTyrIleTyrAspIleGlyAsnThrTyrAsnProSerLeuLysSer 66
OY 175 -----ATGGCTTAATCTGATTAACATATCTCAACAGGATGTACAGCGCA 225
    ||| ::::
DB 67 ArgValThrMetSerValAspThrSerLysAsnGlnHisSerLeuGlnLeuThrSerVal 86
OY 226 ACG-----AAGTAAGTACTACTATCCAGTCAAC 255
    ||| ||| ||| |||
DB 87 ThrAlaAlaAspThrAlaValTyrTyrCysAlaArgValProTyrTyrTyrAspThrSer 106
OY 256 GACGGTTTTCCCGGAGTCACTAT-----GCGGTG 285
    ||| ::::
DB 107 GlyGlyTyrLeuGlnGlyGlnTyrTyrTyrGlyMetAspValThrProGlnGlnThrLeuVal 126

```

Page 8

```

Db      1021  ProThSerAlaLysGluMetThrLysLeu-----GlyValThrAlaGluIvtr 1037
          |||  |||||||:|||||  :|||:  |||:
Qy      113  TcTCGGAGAAAGACAGACAGACACCCTCAGCATCGCA-----  75
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      1038  SerProAlaSerAivSerLeuGlyThrSerProSerProLInThrThrValValSerThr 1053
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Qy      74  -----TGGCATATCATATGATCCATTCGCTGTCCTCTCCA 36
          |||:  |||:  |||:  |||:  |||:  |||:  |||:  |||:  |||:  |||:
Db      1058  AlaGluAspLeuAlaProLysSerSerAlaThrPheAlaValGlnSerSerThrGlnSerPro 1077
          |||:  |||:  |||:  |||:  |||:  |||:  |||:  |||:  |||:  |||:
Qy      35  TGAGTGAAGACTCCGACGTTCCGTCGTAAGCT 3
          |||:  |||:  |||:  |||  |||
Db      1078  ThrThrLeuSerSerSerAlaSerValAnsSer 1088

RESULT 15
US-09-924-417-60
? Sequence 60, Application US/09924417
? Patent No. US2002014241A1
? GENERAL INFORMATION:
? APPLICANT: Falb, Dean
? TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
? THE TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR
? DISEASE
? NUMBER OF SEQUENCES: 67
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: PENNIE & EDMONDS LLP
? STREET: 1155 Avenue of the Americas
? CITY: New York
? STATE: NY
? COUNTRY: USA
? ZIP: 10036-2711
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Diskette
? COMPUTER: IBM Compatible
? OPERATING SYSTEM: DOS
? SOFTWARE: PastSEO Version 2.0
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/924,417
? FILING DATE: 07-Aug-2001
? * CLASSIFICATION: <unknown>
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US/09/034,286
? FILING DATE: 04-MAR-1998
? APPLICATION NUMBER: 08/870,434
? FILING DATE: 06-JUN-1997
? APPLICATION NUMBER: 08/799,910
? FILING DATE: 13-FEB-1997
? APPLICATION NUMBER: 60/011,787
? FILING DATE: 16-FEB-1996
? APPLICATION NUMBER: 08/599,654
? FILING DATE: 09-FEB-1996
? APPLICATION NUMBER: 08/485,573
? FILING DATE: 07-JUN-1995
? APPLICATION NUMBER: 08/386,844
? FILING DATE: 10-FEB-1995
? ATTORNEY/AGENT INFORMATION:
? NAME: Coruzzi, Laura A
? REGISTRATION NUMBER: 30,742
? REFERENCE/DOCKET NUMBER: 7853-114-999
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (212)7809090
? TELEFAX: (212)8699741
? TELEX: 66141 PENNIE
? INFORMATION FOR SEQ ID NO: 60:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1481 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: unknown
? MOLECULE TYPE: protein
? SEQUENCE DESCRIPTION: SEQ ID NO: 60:
US-09-924-417-60

```


2088.975 Million cell updates/sec

Title: US-09-980-370-5
 Perfect score: 731
 Sequence: 1 gcagacttcacggaacgt.....ccaaggaagccaatgcygcc 402

```
Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283274 seqs 66134437 residues
```

Total number of hits satisfying chosen parameters: 566448

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

```
Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

Command line parameters:
-MODEL=frame+020 model

[illegible]

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result	Score	Query Match	Length	DB	ID	Description
1	716	97.9	1704	2	A53436	gingipain R (EC 3
2	715	97.8	1732	2	T30889	lysine-specific cy
3	711	97.6	2828	2	T28551	hemagglutinin A -
4	706	97.6	1576	2	S47035	gingipain R (EC 3
5	90	92.4	286	2	D75217	probable 2-acetyl
6	90	92.4	286	2	D75217	hypothetical protei
7	89	92.2	418	2	T00124	hypothetical protei
8	88.5	92.1	218	2	C81937	probable hydrolase
9	88.5	92.1	218	2	C81937	cellulase (EC 3.2
10	88.5	92.1	428	2	S03722	probable hydrolase
11	87.9	91.4	373	2	B70819	proteobacterial cel
12	82.5	91.3	536	2	A92823	hypothetical protei
13	81.5	91.1	966	2	G65199	hypothetical protei
14	80	90.9	759	2	H84451	cycloartenol synth

[illegible]

ALIGNMENTS

thermolysin (EC 3.
neutral proteinase
alcalase (EC 3.2.
trypsin-like protease)
cellulase (EC 3.2.
genome polyprotein
viral protein - Sh1
microbial metalloprotease
alkaline amylose (EC
3.2.1.6)
probable acid-CoA
endoxylanase - Tum
integrin alpha-V C
cyclo-oxygenant synth
phosphatidylesterase
bacteriolysin (EC 3.
probable serine/thr
tryptophanase (EC 3.2.
cellulase (EC 3.2.
hypothetical protease
hypothetical protease
hypothetical prote
probable sea prot
microbial metalloprotease
neural proteinase
chitinase (EC 3.2.
thermolysin (EC 3.
beta-glucosidase-type
hypothetical prote
alpha-amylose (EC

RESULT 1
255426

glinglabin R (accession number: A55426). Precursor - Porphyromonas gingivalis
N: Alternate names: 50K high molecular mass arginine-specific cysteine protease; HGP
C: Species: Porphyromonas gingivalis
C: Date: 10-Feb-1995 #sequence, revision 10-Feb-1995 #text-change 08-Oct-1999

A: Accession: A55426; D53113 pI: 6.7 N: N.; Prochazka, V.; Klefer, M.C.; Trevis, J.; Barrett,
T.J.; Travis, J. ; PMID:95130800
A: Title: Molecular cloning and structural characterisation of the Arg-ginglabin protease
A: Reference number: A55426; PMID:95130800; PMID:7836351
A: Accession: A55426
A: Status: preliminary
A: Molecule type: DNA
A: Chemical formula: C₁H₁₀N₂O₂S
A: Database accession: GB:U15282; NID:9557067; PID:NAA69539.1; PTD:g557068
R: PI: R.; McGraw, E.; Potempa, J.; Trevis, J.
J: Biol. Chem. 269, 406-411, 1994
A: Title: Lysine- and arginine-specific proteinases from Porphyromonas gingivalis. Iso
#: Reference number: A53113; PMID:94103245; PMID: 8276627

Alignment Scores:

A: NCBI accession: U031137
A: Molecular Weight: 69kDa protein
A: Residues: 328-249 <PK>
A: Experimental source: H66
A: Note: sequence extracted from NCBI backbone (NCBP:141654)
C: Keywords: cysteine proteinase; hydrolase

Pred. No.:	3,926-bb	Length:	170
Score:	716.00	Matches:	170
Percent Similarity:	99.25%	Mismatches:	1
Best Local Similarity:	97.76%	Mismatches:	2
Query Match:	97.95%	Indels:	0
DB:	2	Gaps:	0

OY 1 GCAGACTTCACGGAACGCTTCGAGCTTCTACTGATGAGAGGACGACGCAATGGACT 60
 |||
 Db 1137 Alasppherthrgluthrphgglusersthrhsglualaprolaglutprphr 1156
 OY 61 ACTATGATGCCGATGCCGATGCGAGGCTGGCTGTCTCTCCGACCAATTGGAC 120
 |||
 Db 1157 ThrileaspAlasppolyspdlgylindlytrpleucylseuserstgylindlasp 1176
 OY 121 TGGCTCAGACGCTGAGGCGACCAACGATGAGCACTTCTCATGCAATGAGTGGCT 180
 |||
 Db 1177 TrpleuthrAlahlsclgylthrasnvalvalalaserphesertrpansglymelala 1196
 OY 181 TTGATCTGCTGATCTCATCTCAAGGATGACGCGACCAAGCAATGAAAGTAC 240
 |||
 Db 1197 LeuasnproaspasntylleuileserLysaspvalthrglualathrLysvallystyr 1216
 OY 241 TACTATCCAGTCAACGAGCTTTCGCGGATCAGTACGCTGATGATCTCCAAAGC 300
 |||
 Db 1217 TytrAlaValalasnspgllyserthrglualasnglyala 1236
 OY 301 GCGACGACGCGGACGACTTCACGAGTGTTCGCAAGAACGCTACGCAATGAATAG 360
 |||
 Db 1237 GylthrasnAlahlsclgylasppherthvalalpheglnlutlthrProasnlylileasnlys 1256
 OY 361 GCGGACGCAACGATTCGCTTCACGCAAGCAATGGCGCC 402
 |||
 Db 1257 GylgylAlahlsrphggllyleuaserthrglualasnglyala 1270

RESULT 2
 T30836
 Lysine-specific cysteine proteinase porphyraia (EC 3.4.22.-) - Porphyromonas gingivalis
 N:Alternate names: Lysine-specific cysteine proteinase 1, 60K
 C:Species: Porphyromonas gingivalis
 C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 17-Nov-2000
 A:Accession: T30836; T30837; T30526; A53113
 J:Bio1. Chem. 269, 406-411, 1994
 A:Title: Lysine- and arginine-specific proteinases from Porphyromonas gingivalis. Isolat
 A:Reference number: A53113; NUID:94103245; PMID:8276827
 A:Accession: A53113
 A:Molecule type: DNA
 A:Status: preliminary
 A:Notes: sequence extracted from NCBI backbone (NCBP:141690)
 A:Gene: prtB, prtK

C:Keywords: cysteine proteinase; hydrolase
 Alignment Scores:
 Pred. No.: 5,06e-68 Length: 1732
 Score: 715.00 Matches: 131
 Percent Similarity: 97.258 Conservative: 2
 Query Match: 97.814 Gaps: 0
 DB: 2 Gaps: 0

US-09-980-370-5 (1-402) x T30836 (1-1732)
 OY 1 GCAGACTTCACGGAACGCTTCGAGCTTCTACTGATGAGAGGACGACGCAATGGACT 60
 |||
 Db 1157 Alasppherthrgluthrphgglusersthrhsglualaprolaglutprphr 1176
 OY 61 ACTATGATGCCGATGCCGATGCGAGGCTGGCTGTCTCTCCGACCAATTGGAC 120
 |||
 Db 1177 ThrileaspAlasppolyspdlgylindlytrpleucylseuserstgylindlasp 1196
 OY 121 TGGCTCAGACGCTGAGGCGACCAACGATGAGCACTTCTCATGCAATGAGTGGCT 180
 |||
 Db 1197 TrpleuthrAlahlsclgylthrasnvalvalalaserphesertrpansglymelala 1216
 OY 181 TTGATCTGCTGATCTCATCTCAAGGATGACGCGACCAAGCAATGAAAGTAC 240
 |||
 Db 1217 TytrAlaValalasnspgllyserthrglualasnglyala 1236
 OY 241 TACTATCCAGTCAACGAGCTTTCGCGGATCAGTACGCTGATGATCTCCAAAGC 300
 |||
 Db 1237 GylthrasnAlahlsclgylasppherthvalalpheglnlutlthrProasnlylileasnlys 1256
 OY 301 GCGACGACGCGGACGACTTCACGAGTGTTCGCAAGAACGCTACGCAATGAATAG 360
 |||
 Db 1257 GylthrasnAlahlsclgylasppherthvalalpheglnlutlthrProasnlylileasnlys 1276
 OY 361 GCGGACGCAACGATTCGCTTCACGCAAGCAATGGCGCC 402
 |||
 Db 1277 GylgylAlahlsrphggllyleuaserthrglualasnglyala 1290

RESULT 3
 T28651
 Degradinglutin A - Porphyromonas gingivalis
 C:Species: Porphyromonas gingivalis
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 03-Aug-2001
 A:Accession: T28651
 R:Han, N.; Whitlock, J.; Proguiske-Fox, A.
 Infect. Immun. 64, 4000-4007, 1996
 A:Title: The hemagglutinin gene A (hgaA) of Porphyromonas gingivalis 381 contains fou
 A:Accession: T28651; Z20434; NUID:97047672; PMID:8926061
 A:Status: preliminary
 A:Molecule type: DNA
 A:Accession: T28651
 A:Cross-references: EMBL:U41807; NID:g1552410; PID:g1469916; PIDN:AA817128.1
 A:Gene: hgaA

Alignment Scores:
 Pred. No.: 1.42e-67 Length: 2628
 Score: 711.00 Matches: 130
 Percent Similarity: 97.258 Conservative: 3
 Query Match: 97.264 Gaps: 0
 DB: 2 Gaps: 0

US-09-980-370-5 (1-402) x T28651 (1-2628)
 OY 1 GCAGACTTCACGGAACGCTTCGAGCTTCTACTGATGAGAGGACGACGCAATGGACT 60
 |||
 Db 685 Alasppherthrgluthrphgglusersthrhsglualaprolaglutprphr 704
 OY 61 ACTATGATGCCGATGCCGATGCGAGGCTGGCTGTCTCTCCGACCAATTGGAC 120

D75217
 Probable 2-acetyl-1-nitroglucylglycerophosphoethanolamine esterase PAB2176 - *Pyrococcus abyssi* (Strain
 C:Species: *Pyrococcus abyssi*
 C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
 C:Accession: D75217
 R:Anonymous, Genoscope
 Submitted to the EMBL Data Library, July 1999
 A:Description: *Pyrococcus abyssi* genome sequence: Insights into archaeal chromosome structure
 A:Reference number: A75001
 A:Accession: D75217
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-286 <RAN>
 A:Cross-references: GB:AJ248283; GB:AL096836; NID:95457433; PIDN:CA9187.1; PID:ea151508
 A:Experimental source: strain Orsay
 C:Genetics:
 A:Gene: PAB2176

Alignment Scores:
 Pred. NO.: 0.232 Length: 286
 Score: 90.50 Matches: 29
 Percent Similarity: 43.27% Conservative: 16
 Best Local Similarity: 27.88% Mismatches: 26
 Query Match: 12.38% Indels: 33
 2 Gaps: 5

US-09-980-370-5 (1-402) x D75217 (1-286)

OY	16	ACGTCGAGCTTCTTACTCATGAGAGGACACGCGAATGACCTCATGATGCCGAT	75
DB	102	ThPhhAaPpHeArGAlaHhScIyLubSeRluClySeRlyStrPhrIle-----	118
OY	76	GCGGATGCGAGGCGTTCGCTGTCTGCTTCGCGACAAATGGACGTGGCCACAGTCAT	135
DB	119	GIYAspIySglu-----IleLuuSpIeuSeRlYAlaIleAspIrlleuSeuSeRn	136
OY	136	GCGGCGACCAAGCTAGTAACTCTTCTCATGCAAT-----	171
DB	137	ThAsnThrLysIleAlaLeuIleGlyHeSeRreTolYAlaMetAlThrIleArg	156
OY	172	-----GGAATGCGTTTGAACTCGTAACTATATC	201
DB	157	AlaLeuAlaGluAspIuArYAlCysCsgIYleAlaAspErlProIleYrIle	176
OY	202	ATCTCAAGAGTTGTACAGCGCCACGAGGTAAGTAACTATAC-----	243
DB	177	-----AspIyStrHcIYAlaArgIYleuIyStrYrHeAlaAsnLeuProGluPhe	193
OY	244	---TATCCAGTC	252
DB	194	LeuYrProIle	197

RESULT 7
 T000154
 Hypothetical protein 21 - *Staphylococcus aureus* phage phi PVL
 C:Species: *Staphylococcus aureus* phage phi PVL
 C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 21-Jul-2000
 C:Accession: T000154
 R:Kameko, J.; Kimura, T.; Kawakami, Y.; Tomita, T.; Kameo, Y.
 Biosci. Biotechnol. Biochem. 61, 1960-1962, 1997
 A:Title: Pantoon-Valettine leukocidin genes in a phage-like particle isolated from mitomy
 A:Reference number: Z14119; M0ID:98067870; PMID:9404084
 A:Accession: T000154
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-418 <RAN>
 A:Cross-references: EMBL:AB009866; NID:93341907; PIDN:BA31894.1; PID:93341928
 C:Superfamily: *Staphylococcus aureus* phage phi PVL hypothetical protein 21

Alignment Scores:
 Pred. NO.: 0.35 Length: 418
 Score: 89.00 Matches: 36
 Percent Similarity: 41.67% Conservative: 19

```

best local similarity: 27.27%      mismatches: 55
Query Match: 12.18%              indels: 22
DB: 2                             gaps: 7

US-09-980-370-5 (1-402) x T00154 (1-418)

OY 58 ACTACTATGATGCCGATGGCGATGTGAGCGTTGCCTGTCTCTTCGACAAATTG 117
    ||| ||||| :||||| |
Db 94 ThTtYrLleAspGlYgLygluaSpolYcyl-----SerSerlyThrIle 109
OY 118 GATCGT-----CTCACAGCCTATGGCCGCCAACAGCTACTAAGCTCTTTCATG 168
    ||| ||||| :||| ||| ||||| :|||
Db 110 GlHTrPTraSPluSThrtTyrsSerAspScRlyMeLsnclYlLeHrlLnsSerlyr 129
OY 169 AATGA---ATGGCTTTGCAATCCTGATTAATCTCATCTCAAAGATTTACAGCCCA 225
    ||| ||||| :||| ||| :|||
Db 130 GLYGLYValAlaLalenuThSerAspAsnArYValLleuLnuSerlyrAlaser 149
OY 226 ACAAGATAAG-----TACTACTATCCAGTCAACGACGCTTTTCCC 267
    ||| ||||| :||| ||||| ||| |||
Db 150 SerAsnLleYserSerlyGlnAlaProValTYrLentyrProAsnthrasplysValPro 169
OY 268 GGG---GATCGACATGGCGGTGATGATCTCCAAAGC---GGCAGCAAGCCGGAGACATTC 321
    ||| ||||| :||||| :|||||
Db 170 GLYLeuAsnArYphenAlaPhethrLenuSerAsnAlaAspAnAlarySerSerAspely 189
OY 322 AGCGTTGTTTTGCAAGAAGCCCTAACCGAATAATAAATGAAGCGCGACAGATTC----- 375
    ||||| ||| ||| |||||
Db 190 TYrILeMetPheGlySerAspPoluAsnArYrAspyrIoLYalacLYlLeArgPheSerlyS 209
OY 376 -----GRTCTTTCACGAGCAACCAGATGC 399
    ||||| |||||
Db 210 GLuArgAsnLySGlyLeuValGlnLeValasncly 221

RESULT 8
C81847
Probable hydrolase NMA1552 [Imported] - Neisseria meningitidis (strain Z2491 serogroup O:3Specie: Neisseria meningitidis
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
A:Accession: C81847
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mo
J.: Holroyd, S.; Jørgens, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandre
Nucleotide 404, 502-506, 2000
A>Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491
A:Reference number: AB1775; MUID:20222556; PMID:10761919
A:Accession: C81847
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-218 <PAR>
A:Cross-references: GB:AL162756; GB:AL157959; MID:g7380091; PIDN:CAB84779.1; PID:g738
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: NMA1552
C:Superfamily: 2-hydroxyhepta-2,4-diene-1,7-dioate isomerase

Alignment Scores:
Pred. No.: 0.369 Length: 218
Score: 88.50 Matches: 38
Percent Similarity: 36.36% Conservative: 14
Best Local Similarity: 26.57% Mismatches: 48
Query Match: 12.11% Indels: 43
DB: 2 Gaps: 6

US-09-980-370-5 (1-402) x C81847 (1-218)

OY 67 GATGCCGATGGCGATGTGAGCGT----- 90
    |||:||||| |||||
Db 77 AspSerAspGlyThrclyGluGlyLysAspIleuGlyCysValAlaIaclyVal 96
OY 91 -----TGCGCTGTCGTCTTCGCGACATTTGACTCG 123
    ||||| ||| ||| |||
Db 97 GlyLeuAspLeuThrAlaArgAspIleGlnCysThrLeuLysGlyLysGlyLeuProTrp 116
```

OY 124 CTCACAGCTCATGGC-----GGCACCACAGTACTAAGCTCTTTTCATGGAATGCAATG 177
DB 117 LeuLysAlaLysSclYpheaRgHnIsSerAlaCyValSerAspPheaAlaAlaGlyArg 136
OY 178 GCTTGAATCTGATATACTATCATCTCAAGAGATGTTACAGCGCAGACAGCAAGTAAAG 237
DB 137 IleGlyAsnProGlnAsnValLeuPheSerLeuLysGlnAsnGlyValLeuLysGlnArg 156
OY 238 -----TACTACTATCCAGTCACAGCGGTTTCCCGGGATCATATGCG 282
DB 157 GlyAspThrGlyLeuMetIleTyrProIleArgGlu----- 168
OY 283 GTGATGATCTCCCAAGAG-----GGCAGCAGACCCGAGACTTCACGCTTGT 330
DB 169 -----IleLeuHisLysLeuAlaAlaAspTyrGlyLeuGlyLysGlyAsp-----LeuVal 185
OY 331 TTCGAGAAACCGCTTACAGCAATTAATTAAGCGCGCAGCAAGATTCGCTTTCACAGGAA 390
DB 186 PheThrGlyThrSerSerGlyValGlyAlaIleGlyAlaGlyAspAsnLeuAlaLeuGlu 205
OY 391 GCCAATGCC 399
DB 206 LeuAspGly 208

RESULT 9

S03767
cellulase (EC 3.2.1.4) 2 precursor - Erwiniia chrysanthemi
N:Alternate names: endo-1,4-beta-glucanase; extracellular endoglucanase 2
C:Species: Erwiniia chrysanthemi
A:Variety: strain 3937
C>Date: 21-Nov-1993 #sequence_revision 01-Dec-1995 #text_change 08-Oct-1999
R:Gutsepri, A.; Cami, B.; Aymeric, J.L.; Ball, G.; Creuzet, N.
Mol. Microbiol. 2, 159-164, 1988
A:Title: Homology between endoglucanase 2 of Erwiniia chrysanthemi and endoglucanases of
A:Reference number: S03767; MUID:88216177; PMID:2835589
A:Accession: S03767
A:Molecule type: DNA
A:Residues: 1-428 <GUI>
A:Cross-references: EMBL:Y00540; NID:g41091; PIDN:CAA68604.1; PID:g41092
C:Genetics:
A:Gene: celZ
C:Function:
A:Description: hydrolysis of 1,4-beta-D-glucosidic linkages in beta-D-glucans such as ce
A:Pathway: cellulose degradation
C:Keywords: glycosidase; hydrolase; polysaccharide degradation
F:1-43/Domain: signal sequence #status predicted <SIG>
F:44-428/Product: cellulase 2 #status predicted <MAT>

Alignment Scores:

Pred. No.:	1.06	Length:	428
Score:	84.50	Matches:	37
Percent Similarity:	40.74%	Conservative:	18
Best Local Similarity:	27.41%	Mismatches:	30
Query Match:	11.56%	Indels:	50
DB:	2	Gaps:	7

US-09-980-370-5 (1-402) x S03767 (1-428)

OY 7 TTCACGGAAAGCTTGAAGTCTTCTACTATGAGAGCA----- 45
DB 232 TTTTLeuHisRheTyrAlaGlyThrHisGlyLysSerLeuAlaGlnLysAlaArgIn 251
OY 46 -----CCAGCGAATGAGTACTATGATCGATCGGATGCG 78
DB 252 AlaLeuAsnAsnGlyIleAlaLeuRheValThrGluTyrGlyThrValAsnAlaAspGly 271
OY 79 GATGTCAGAGGTGGCTGTCTGTCTCCGAGCAATGAGAC---TGCTCAGCTAT 135
DB 272 AsnGlyGly-----ValAsnGlnThrGluThrAspAlaThrValThrPheMet 287
OY 136 GCGGCGACCAAGCTAGACCTTTTTCATGGAATGGAATGCTTTGATCTGATAC 195

DB 288 ArgAspAsnAsnIleGlnLeuThrGlnAsnTrp-----AlaLeuAsnAspLysAsn 304
OY 196 -----TATCATCTCAAGAGATGTTACA---GGCGCAGCAGGTA 234
DB 305 GluGlyAlaSerThrTyrTyrProAspSerLysAsnLeuThrGluSerGlyLysVal 324
OY 235 AACTACTACTATCCAGTCA----- 253
DB 325 Lys-SerIleIleGlnSerTrpProTyrLysAlaGlySerAlaAlaSerAlaThrThrAs 344
OY 254 -----ACGACGTTTTCGCGGATCATATGCGG 283
DB 344 ProSerThrAspThrMetThrProProLeuThrAsnArg 358

RESULT 10

A46564
thermolysin (EC 3.4.24.27) homolog precursor - Bacillus stearothermophilus
C:Species: Bacillus stearothermophilus
C>Date: 03-Mar-1994 #sequence_revision 03-Mar-1994 #text_change 22-Jun-1999
C:Accession: A46564
R:Kubo, M.; Imanaka, T.
J. Gen. Microbiol. 134, 1883-1892, 1988
A:Title: Cloning and nucleotide sequence of the highly thermostable neutral protease
A:Reference number: A46564; MUID:89235633; PMID:3149972
A:Accession: A46564
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-552 <KUB>
A:Cross-references: GB:M21663; NID:g143746; PIDN:AAB02774.1; PID:g143747
A>Note: the authors translated the codon TCG for residue 32 as Thr, CTC for residue 5
C:Superfamily: thermolysin
C:Keywords: hydrolase; metalloproteinase; zinc
F:378,382,402/Binding site: zinc (His, His, Glu) #status predicted
F:379,467/Active site: Glu, His #status predicted

Alignment Scores:

Pred. No.:	1.4	Length:	552
Score:	83.50	Matches:	41
Percent Similarity:	42.59%	Conservative:	28
Best Local Similarity:	25.31%	Mismatches:	52
Query Match:	11.42%	Indels:	41
DB:	2	Gaps:	10

US-09-980-370-5 (1-402) x A46564 (1-552)

OY 25 TTTTCTACTCAT-----GGAGAGGCACAGCGGATGAGCT-----ACTATC 66
DB 338 SerSerValHisTyrSerGlnGlyTyrAsnAsnAlaPheThrAsnGlySerGlnMetVal 357
OY 67 GATCCGATGCGCGAGTGGTGGCTGTCTGTCTCCGAGCAATGAGCTGGCTC 126
DB 358 TyrGlyAspGlyAspGlyGlnThrPheIleProLeuSerGlyGly---IleAsp---Val 375
OY 127 ACAGTCATGCGCGCAGCAGCAAGTAAAGCTTTTCTCATGATGAATGCTTTGAAT 186
DB 376 ValAlaHisGluLeuThrHisAlaValThrAspTyrThr---AlaGlyLeuIleTyrGln 394
OY 187 CCTGATACATCTCATCTCAAG-----GATGTTACAGCGCAGCAGCAAGTAAAG 237
DB 395 AsnGluSerGlyAlaIleAsnGlnAlaIleSerAspIlePheGly---ThrLeuValGlu 413
OY 238 TACTACTATTCAGTCAAGCGGTTT-----CCGCGG--- 270
DB 414 PheTyrAlaAsnLysAsnProAspTyrGluIleGlyLysAspValTyrThrProGlyIle 433
OY 271 -----GATCAGTAT 279
DB 434 SerGlyAspSerLeuArgSerMetSerAspProAlaLysTyrGlyAspProAsnHisTyr 453
OY 280 GCGGTGATGATCTCCAGAGCGGCGCAGACCGCGAGACTTC-----ACGGTTGTTTC 333
DB 454 SerLysArgTyrThrGlyThrGlnAspAsnAlaGlyValHisIleAsnSerGlyIleIle 473

OY 334 GAAGAACGCGTAAACGGAATAAATAGCGCGGACGAAGATTGCGCTTTCCACGGAGCC 393
Db 474 AsnlysaAlaAlaTyrleuileSerGInGlyGlyThrIstYrGlyAlaSerValaGly 493
OY 394 AATGGC 399
Db 494 IleGly 495

RESULT 11
B70819
probable carnitine operon oxidoreductase - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: B70819
R:COLE, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634220
A:Accession: B70819
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-373 <COL>
A:Cross-references: GB:AL022002; GB:AL123456; NID:q3261544; PIDN:CAAI7610.1; PID:q326154
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: fadE16
C:Superfamily: acyl-CoA dehydrogenase

Alignment Scores:
Pred. No.: 1.5E Length: 373
Score: 83.00 Matches: 39
Percent Similarity: 39.13% Conservative: 15
Best Local Similarity: 28.26% Mismatches: 64
Query Match: 11.35% Indels: 20
Gaps: 2 6

US-09-980-370-5 (1-402) x B70819 (1-373)

OY 25 TCTTCTACTCATGAGACGACGACGAGATGACTATCGATGCGGATGGCATGCT 84
Db 118 SerATgSerHIsheTrrpAlaProVal-----SerHrAlaSerAlaAspGlyAspGly 135
OY 85 GAGGGTGGCTGCTGCTGCTTCCGACAAATTGAGCTGCTCACAGCTCATGGCGGAC 144
Db 136 -----IleAlaValAlaGAlaAspLysSerTrrpValThrSerAlaGlyPheAla 151
OY 145 AAC-----GTAGTAGCTCTTTCATGATGATGATGATGATGATGATGATGATGAT 198
Db 152 AspValTrrpValAlaSerValGlySerAlaAspGlyAlaAlaGlyAspValAspLys 171
OY 199 CTCATCTCAAGAGGATTTACAGCGCGCAACGAGTAAGTAAAGTACTATCCAGTCAAGC 258
Db 172 AlaValAlaProAlaAspHrProGlyLeuArgValAlaGlyThrPhe-----ThGlyMet 189
OY 259 GGTTCCTCCGGGATCAGTACATGCGGATGATC-----TCC 294
Db 190 GlyLeuArgGlyAsnAlaSerAlaProMetAlaValAlaSerAlaArgIleProAspSerTyr 209
OY 295 AAGAGCGGACGACGACGAGACTTCAGGTTGTTTCGAGAAAGCCCTAAGCAATA 354
Db 210 ArgLeuGlyGluAlaGlyGlyGlyPheGlyIleMetLeuGlnTrrpValLeuProTrrpPhe 229
OY 355 AATAGCGGCGACCA-----AGATTTCGCTTTCCACGAGCAATGGCGCC 402
Db 230 AsnLeuGlyAsnAlaAlaValSerLeuGlyLeuAlaThrAlaIleThrGlyAla 247

RESULT 12
A99283
hypothetical protein dppC-2 [imported] - Sulfolobus solfataricus
C:Species: Sulfolobus solfataricus

C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001
C:Accession: A99283
R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Ch
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder
arrett, R.A.; Ragan, M.A.; Sengen, C.W.; Van der Oost, J.
submitted to Genbank, April 2001
A:Description: Sulfolobus solfataricus complete genome.
A:Reference number: A99139
A:Accession: A99283
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-536 <KUN>
A:Cross-references: GB:AE006641; NID:q13814480; PIDN:AAK41520.1; GSPDB:GN00155
C:Genetics:
A:Gene: dppC-2

Alignment Scores:
Pred. No.: 1.78 Length: 536
Score: 82.50 Matches: 41
Percent Similarity: 38.89% Conservative: 22
Best Local Similarity: 25.31% Mismatches: 34
Query Match: 11.29% Indels: 65
Gaps: 2 10

US-09-980-370-5 (1-402) x A99283 (1-536)

OY 76 GCGATGGTACGCGTTCGCTGCTGCTTCCGACAAATTGAGTGGCTCACAGCTCAT 135
Db 164 GlyAspTrrpGluPheTrrpIleIleGlnAlaSer-----IleSerAlaTrrp 178
OY 136 GCGGCGACCAAGCTGTAGTACCTTTTCATGATGATGATGATGATGATGATGATGATGAT 192
Db 179 ---AlaThrAsn-----AlaTrrpTrrpPheAsnGlyTrrpValIleAsnProSerAsn 195
OY 193 -----AACTATCTCATCTCAAGAGATGTTACAGCGCA 225
Db 196 AlaThrPheTrrpLeuPheLeuAlaGlyAsnTrrpLeuProThrAsnLeuValThrLeuSer 215
OY 226 ACGAAGTAAATAC----- 240
Db 216 ThrValPheLysTrrpLeuGlyAsnGlyTrrpAsnTrrpIleLeuAlaSerAlaSer 235
OY 241 -----TACTATCCAGTCAACGAGCGGTTTCCGGGATCATATGCG 282
Db 236 AlaGlyGluThrProTrrpPheTrrpThrAsn-----IleProProAsnGluSerAla 253
OY 283 GTG-----ATGATCTCCAAGACGGGACGACGACGCGGAC 318
Db 254 ValAlaSerValIleMetLeuGlnSerMetPheAsnSerThrGlyAsnTrrpLysValGlu 273
OY 319 TTCACGGTGTTCGACAAACGCGCTAAGCGAATAAAGGCG----- 363
Db 274 PheThrIleAsnTrrp-----IleProAsnGlyProAsnSerIleuValIleTrrpLeu 291
OY 364 -----GGAGCAAGATTGCGT-----CTTCCACGAGCAAT 396
Db 292 SerAspLeuTrrpPheGluPheLeuGlySerArgTrrpGlyValLeuGlyThrAspAsn 311
OY 397 GCGGCC 402
Db 312 GlyAla 313

RESULT 13
G69189
hypothetical protein MTH674 - Methanobacterium thermoautotrophicum (strain Delta H)
C:Species: Methanobacterium thermoautotrophicum
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
C:Accession: G69189
R:Smith, D.R.; Doucet-Stamm, L.A.; Delonghery, C.; Lee, H.; Dubois, J.; Aldredge, T
; Qiu, D.; Spadafora, R.; Viciore, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani,
K.I.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: fu

A:Reference number: A69000; MUID:98037514; PMID:9371463
 A:Accession: G69189
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-966 <MTH>
 A:Cross-references: GB:AE000847; GB:AE000666; NID:g2621756; PIDN:AA865179.1; PID:g262175
 A:Experimental source: strain Delta H
 C:Genetics:
 A:Gene: MTH674

Alignment Scores:
 Pred. No.: 2.43 Length: 966
 Score: 81.50 Matches: 26
 Percent Similarity: 50.006 Conservative: 10
 Best Local Similarity: 36.118 Mismatches: 29
 Query Match: 11.158 Indels: 7
 DB: 2 Gaps: 3

US-09-980-370-5 (1-402) x G69189 (1-966)
 Oy 202 ATCTCAAGAGATTACAGCGCGCAAGTAAGTACTATTCAGTCTC----- 252
 Db 296 ILeSeGlyTyrAlaMetSerLeuThrLeuLysTyrTyrProProValMetAspPhe 315
 Oy 253 ---AACGACGGCTTTCCGGCGGCGGATCATATGCG---GTGATGATCTCCAAAGCGGCGACG 306
 Db 316 ThrAsnProGlySerProGlyGluIleThrSerTyrValThrValGlyValProGlyAsp 335
 Oy 307 AACGCGGAGACTTCACGCGTTGTTTCGAAGAACGCTTACGCGAATAATAGCGCGCA 366
 Db 336 SerAspAspAspSerLeuMetPheLeuAspThrThrProGly-----LysGlyGly 353
 Oy 367 GCAAGATGGCTCTTCACGCGAAGCAATGGCGCGC 402
 Db 354 SerTyrSerGlyPheAsnThrThrAlaThrGlyAla 365

RESULT 14
 HB4481
 cycloartenol synthase [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
 C:Accession: H84481
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
 Euser, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
 Nature 402:761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: A84420; MUID:20083487; PMID:10617197
 A:Accession: H84481
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-759 <STO>
 A:Cross-references: GB:AE002093; NID:g6598481; PIDN:AAC69112.2; GSPDB:GN00139
 C:Genetics:
 A:Gene: Atg907050
 A:Map position: 2
 C:Superfamily: yeast lanosterol synthase

Alignment Scores:
 Pred. No.: 3.42 Length: 759
 Score: 80.00 Matches: 25
 Percent Similarity: 43.688 Conservative: 13
 Best Local Similarity: 28.748 Mismatches: 15
 Query Match: 10.948 Indels: 34
 DB: 2 Gaps: 7

US-09-980-370-5 (1-402) x H84481 (1-759)
 Oy 37 GGAGAGCACCACCGAATGACTACTATGATCGGATGCGATGAGGTTGGCTC 96
 Db 188 GlyGluGlyPro-----AsnAspGlyAspGlyAsp----- 197
 Oy 97 TGTCGTCTTCGCGGACCAATGATGATGCGTCAATGCGGCGC---ACCAACGTACTGA 153

Db 198 ---MetLeuLysGlyArg---AspTrpIleLeuAsnHisGlyAlaThrAsnIleThr 215
 Oy 154 AGC-----TCTTCTCATGGAATGCA---ATGCGT 180
 Db 216 SerTrpIleLysMetTrpLeuSerValLeuGlyAlaPheGluTrpSerGlyAsnAsnPro 235
 Oy 181 TTGATCGCTGATATCATATCATCTCAAGAGATGTTACAGCGGCAACGAGTAAGTAC 240
 Db 236 LeuProProGluIleTrpLeuLeu-----ProTyr 245
 Oy 241 TACTATCAGCTCAACGACGCT 261
 Db 246 PheLeuProIleHisProGly 252

RESULT 15
 HB57
 thermolysin (EC 3.4.24.27) precursor [validated] - Bacillus "thermoproteolyticus"
 N:Alternate names: Bacillus thermoproteolyticus neutral proteinase; microbial metallo
 C:Species: Bacillus "thermoproteolyticus"
 C:Date: 24-Apr-1984 #sequence_revision 07-Jun-1996 #text_change 15-Sep-2000
 C:Accession: I40579; A00993; S41312
 R:O'Donohue, M.J.; Roques, B.P.; Beaumont, A.
 Biochem. J. 300, 599-603, 1994
 A:Title: Cloning and expression in Bacillus subtilis of the npr gene from Bacillus th
 A:Reference number: I40579; MUID:94271180; PMID:8002967
 A:Accession: I40579
 A:Molecule type: DNA
 A:Residues: 1-408, 'K', 410-548 <CON>
 A:Cross-references: EMBL:X76988; NID:g441266; PIDN:CAA54291.1; PID:g441267
 A:Note: Submitted to the EMBL Data Library, December 1993
 R:Litani, K.; Hermodson, M.A.; Ericsson, L.H.; Walsh, K.A.; Neurath, H.
 Nature New Biol. 238, 35-37, 1972
 A:Title: Amino-acid sequence of thermolysin.
 A:Reference number: A00993
 A:Accession: A00993
 A:Molecule type: protein
 A:Residues: 233-268, 'D', 270-350, 'E', 352-399, 'I', 401-548 <TTR>
 R:Matthews, B.W.; Holmes, M.A.
 Submitted to the Brookhaven Protein Data Bank, February 1982
 A:Reference number: A50631; PDB:3FLN
 A:Contents: annotation; X-ray crystallography, 1.6 angstroms, residues 233-268, 'D', 27
 R:Holmes, M.A.; Matthews, B.W.
 J. Mol. Biol. 160, 623-639, 1982
 A:Title: Structure of thermolysin redefined at 1.6 A resolution.
 A:Reference number: A92888; MUID:83085812; PMID:7175940
 A:Contents: annotation; X-ray crystallography, 1.6 angstroms
 R:Matthews, B.W.; Weaver, L.H.; Kester, W.R.
 J. Biol. Chem. 249, 8030-8044, 1974
 A:Title: The conformation of thermolysin.
 A:Reference number: A92165; MUID:75041142; PMID:4214815
 A:Contents: annotation; X-ray crystallography, 2.3 angstroms
 R:Biochemist, Y.; Walsh, K.A.; Neurath, H.
 Biochemistry 13, 205-210, 1974
 A:Title: Evidence of an essential histidine residue in thermolysin.
 A:Reference number: A90377; MUID:74052951; PMID:4808703
 A:Contents: annotation; active site
 C:Function:
 A:Description: metalloendopeptidase preferentially hydrolyzes peptide bonds on the am
 C:Superfamily: thermolysin
 C:Keywords: calcium; extracellular protein; hydrolase; metalloproteinase; zinc
 F:1-18/Domain: signal sequence #status predicted <SIG>
 F:19-232/Domain: propeptide #status predicted <PRO>
 F:233-548/Product: thermolysin #status experimental <CAT>
 F:289,291,293/Binding site: calcium (Asp, Asp, Glu) #status experimental
 F:370,409,417,419,422/Binding site: calcium (Asp, Glu, Asp, Glu, Glu) #status experim
 F:374,378,398/Binding site: zinc (His, His, Glu) #status experimental
 F:375,463/Active site: Glu, His #status predicted
 F:409,415,417,422/Binding site: calcium (Glu, Asn, Asp, Glu) #status experimental
 F:425,426,429,432/Binding site: calcium (Tyr, Thr, Ile, Asp) #status experimental

Alignment Scores:
 Pred. No.: 3.73 Length: 548

Score: 79.50 Matches: 40
Percent Similarity: 41.98% Conservative: 28
Best Local Similarity: 24.69% Mismatches: 53
Query Match: 10.88% Indels: 41
DB: 1 Gaps: 10

US-09-980-370-5 (1-402) x HYST (1-548)

```
OY 25 TCTTCTACTCAT-----GGAGAGGCACGCGGANTGGACT-----ACTATC 66
    ||||| ||| ||| ||| |||
DB 334 SerSerValHisTyrSerGlnGlyTyrAsnAsnAlaPheTrpAsnGlySerGlnMetVal 353
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 67 GATGCCGATGGCGATGGATGGGTTGGCTGTCTGTCTCCGACAATTGGACTGGCTC 126
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 354 TyrGlyAspGlyAspGlyGlnThrPheIleProLeuSerGlyGly---IleAsp---Val 371
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 127 ACAGCTCATGGCGGACCAACAGTACTAGCTCTTCTCATGGAAATGAGTGGCTTGAAT 186
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 372 ValAlaHisGluLeuThrHisAlaValThrAspTyrThr---AlaGlyLeuIleTyrGln 390
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 187 CCTGTAATCTATCTCATCTCAAG-----GATGTACAGCGGCAACGAAGTAAAG 237
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 391 AsnGluSerGlyAlaIleAsnGluAlaMetSerAspIlePheGly---ThrLeuValGlu 409
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 238 TACTACTATTCAGTCACGACGCGTTT-----CCCGGG--- 270
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 410 PheTyrAlaAsnLysAsnProAspTrpGluIleGlyGluAspValTyrThrProGlyIle 429
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 271 -----GATCACTAT 279
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 430 SerGlyAspSerLeuArgSerMetSerAspProAlaLysTyrGlyAspProAspHisTyr 449
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 280 GCGGTGATGATCTCCAAAGACGGGACGAACGCCGAGACTTC-----ACGGTTGTTTC 333
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 450 SerLysArgTyrThrGlyThrGlnAspAsnGlyGlyValHisIleAsnSerGlyIleIle 469
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 334 GAAGAAGCCCTTAACGGAATAATAAGCGCGGAGCAAGATCGGCTTCCACGGAAGCC 393
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 470 AsnLysAlaAlaTyrLeuIleSerGlnGlyGlyThrHisTyrGlyValSerValValGly 489
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 394 AATGCC 399
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 490 IleGly 491
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Search completed: June 26, 2003, 00:46:19
Job time : 41 secs

RESULT 2			
HAGA_PORGI			
ID	HAGA_PORGI	STANDARD:	PRT: 2628 AA.
AC	G51845:		
DT	30-MAY-2000 (Rel. 39, Created)		
DT	30-MAY-2000 (Rel. 39, last sequence update)		
DT	16-OCT-2001 (Rel. 40, last annotation update)		
DE	Hemagglutinin A precursor.		
GN	HAGA.		
OS	Porphyromonas gingivalis (Bacteroides gingivalis).		
OC	Bacteria; Bacteroidetes; Bacteroides; Bacteroidales;		
OC	Porphyromonadaceae; Porphyromonas.		
OX	NCBI_TaxID=837;		
RN	[1]		
RP	SOURCE FROM N.A.		
RC	STRAIN=381;		
RC	MEDLINE=87047672; PubMed=8926061;		
RX	Han N., Whitlock J., Froguskie-fox A.;		
RT	"The hemagglutinin gene A (haga) of Porphyromonas gingivalis 381		
RT	contains four large, contiguous, direct repeats.";		
RU	Infect. Immun. 64:4000-4007(1996).		
CC	-1- FUNCTION: AGGLUTININATES ERYTHROCYTES.		
CC	-1- SIMILARITY: BELONGS TO PERITRIN FAMILY C25.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		

CASL_3	CASL_ARATH	STANDARD:	PRT:	759 AA.
AC	p38605: p92967:			
DT	01-OCT-1994 (Rel. 30, Created)			
DT	15-JUN-2002 (Rel. 41, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	cycloartenol synthase (EC 5.4.99.8) (2,3-epoxysqualene--cycloartenol cyclase)			
GN	CAS1 OR AT2G07050 OR T4E14.16.			
OS	Arabidopsis thaliana (Mouse-ear cress).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;			
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsids.			
NCBI_taxid=3702;				

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. Landsberg erecta;
RX MEDLINE=94089686; Pubmed=7505443;
RA Corey E.J., Matsuda S.P.T., Bartel B.;
RT "Isolation of an Arabidopsis thaliana gene encoding cycloartenol
RT synthase by functional expression in a yeast mutant lacking
RT lanosterol synthase by the use of a chromatographic screen.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:11628-11632(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. Columbia;
RX MEDLINE=20083487; Pubmed=10617197;
RA Liu X., Kaul S., Roundley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.J., Konning C.M., Koo H.L.,
RA Moffat K.S., Cronin L.A., Shen M., Pal G., Van Aken S., Umayam L.,
RA Tallon L.J., Gill J.E., Adams M.D., Carrera A.J., Creasy T.H.,
RA Goodman H.M., Somerville C.R., Copenhagen G.P., Preuss D.,
RA Nierman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,
RA Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT thaliana.";
RL Nature 402:761-768(1999).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. Columbia;
RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RT "RIKEN Arabidopsis full length cDNA clones (RAFLs) sequenced by the
RT SSP consortium (Salk/Stanford/PGECC).";
RL Submitted (APR-2002) to the EMBL/Genbank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: (S)-2,3-epoxysqualene -> cycloartenol.
CC -1- SIMILARITY: BELONGS TO THE TERPENE CYCLASE/MUTASE FAMILY.
CC -1- SIMILARITY: CONTAINS 5 PPTB REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U02555; AAC04931.1; -
DR EMBL: AC005171; AAM15015.1; -
DR EMBL: AY094394; AAM19773.1; -
DR InterPro: IPR001330; PrenylTrans.
DR InterPro: IPR002365; Terpene_Synth.
DR Pfam: PF00432; PrenylTrans; 3.
DR PROSITE: PS01074; TERPENE_SYNTASES; 1.
KW Isomerase; Repeat.
FT REPEAT 147 188 PPTB 1.
FT REPEAT 512 557 PPTB 2.
FT REPEAT 589 629 PPTB 3.
FT REPEAT 638 679 PPTB 4.
FT REPEAT 700 741 PPTB 5.
FT CONFLICT 501 501 A -> E (IN REF. 1).
SQ SEQUENCE 759 AA; 86033 MW; DBA75C57BB1F74D CRC64;

```

Alignment Scores:

```

Pred. No.: 1.3 length: 759
Score: 80.00 Matches: 25
Percent Similarity: 43.688 Conservative: 13
Best Local Similarity: 28.748 Mismatches: 15
Query Match: 10.948 Indels: 34
DB: 1 Gaps: 7

```

US-09-980-370-5 (1-402) x CASI_ARATH (1-759)

```

OY 37 GGAGAGCAGCAGCGAATGACTACTGATCGGATGGGATGGAGGTTGGCTC 96
DB 188 GlytGlytGlyPro-----AsnAspGlyAspGlyAsp----- 197

```

```

OY 97 TGCTGCTCCGACAAATTCGCTACAGCTCATGCGCGC---ACCACGTAGTA 153
DB 198 ---MetGlnuysSLyArg---AsprtleLeuasnHsGlyGlyAlaTrnAsnIleThr 215
OY 154 AGC-----TCTTTCATGAGTAAAGCA---ATGGCT 180
DB 216 SerTrpGlyLysMetTrpLeuSerValLeuGlyAlaPheGluTrpSerGlyAsnAsnPro 235
OY 181 TTGAATCTGATATACATATCATCTCAAGATGTTACAGCGCAGCAAGTAAGTAC 240
DB 236 LeuProGluGluIleTrpLeu-----ProTyr 245
OY 241 TACTATCCAGCAACAGCGGT 261
DB 246 PheLeuProIleHisProGly 252

```

RESULT 4

```

THER_BACTH STANDARD: PRT; 316 AA.
AC P00800;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Thermolysin (EC 3.4.24.27).
OS Bacillus thermoproteolyticus.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1427;
RN [1]
RP SEQUENCE.
RA Titani K., Hermodson M.A., Ericsson L.H., Walsh K.A., Neurath H.;
RT "Amino-acid sequence of thermolysin.";
RL Nature New Biol. 238:35-37(1972).
RN [2]
RP ACTIVE SITE.
RX MEDLINE=74052951; Pubmed=4808703;
RA Burselen Y., Walsh K.A., Neurath H.;
RT "Evidence of an essential histidine residue in thermolysin.";
RL Biochemistry 13:205-210(1974).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS).
RX MEDLINE=83085812; Pubmed=7175940;
RA Holmes M.A., Matthews B.W.;
RT "Structure of thermolysin refined at 1.6-A resolution.";
RL J. Mol. Biol. 160:623-639(1982).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
RX MEDLINE=75041142; Pubmed=4214815;
RA Matthews B.W., Weaver L.H., Kester W.R.;
RT "The conformation of thermolysin.";
RL J. Biol. Chem. 249:8030-8044(1974).
RN [5]
RP STRUCTURE BY NMR OF 255-316.
RX MEDLINE=95086079; Pubmed=7993910;
RA Rico M., Jimenez M.A., Gonzalez C., de Fillippis V., Fontana A.;
RT "NMR solution structure of the C-terminal fragment 255-316 of
RT thermolysin: a dimer formed by subunits having the native
RT structure.";
RL Biochemistry 33:14834-14847(1994).
RN [6]
RP STRUCTURE BY NMR OF 205-316.
RX MEDLINE=97452621; Pubmed=9305992;
RA Conjero-Lara F., Gonzalez C., Jimenez M.A., Padmanabhan S.,
RA Mateo P.L., Rico M.;
RT "NMR solution structure of the 205-316 C-terminal fragment of
RT thermolysin. An example of dimerization coupled to partial
RT unfolding.";
RL Biochemistry 36:11975-11983(1997).
CC -1- FUNCTION: THERMOSTABLE EXTRACELLULAR ZINC METALLOPROTEASE.
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Xaa-|-Leu > Xaa-|-Phe.
CC -1- COPACOR: BINDS 1 ZINC ION AND FOUR CALCIUM IONS WHICH
CC CONTRIBUTE TO THE UNUSUAL THERMAL STABILITY OF THIS ENDOPEPTIDASE.
CC TWO ARE CLOSE TOGETHER AND PARTICIPATE IN A NETWORK OF SALT LINKS
CC ALSO INVOLVING ASP-136, GLU-177, LYS-182, ASP-185, GLU-190, AND

```

CC ASP-191. A THIRD INTERACTS WITH ASP-57 AND A FOURTH WITH ASP-200.
 CC GUU AT POSITION 302 APPEARS TO FORM A SALT LINK WITH LYS-262.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M4.
 DR PIR: A00993. HYDST.
 DR PDB: 4TLN: 31-JAN-84.
 DR PDB: 5TLN: 31-JAN-84.
 DR PDB: 7TLN: 31-JAN-84.
 DR PDB: 8TLN: 30-APR-94.
 DR PDB: 1TMN: 09-JAN-89.
 DR PDB: 2TMN: 15-JAN-91.
 DR PDB: 3TMN: 09-JAN-89.
 DR PDB: 4TMN: 09-JAN-89.
 DR PDB: 5TMN: 09-JAN-89.
 DR PDB: 6TMN: 09-JAN-89.
 DR PDB: 7TMN: 15-OCT-94.
 DR PDB: 1THL: 31-DEC-94.
 DR PDB: 1THL: 09-JAN-89.
 DR PDB: 1HTP: 31-JUL-94.
 DR PDB: 1TRL: 07-FEB-95.
 DR PDB: 1LNA: 08-MAY-95.
 DR PDB: 1LNC: 08-MAY-95.
 DR PDB: 1LND: 08-MAY-95.
 DR PDB: 1LNE: 08-MAY-95.
 DR PDB: 1LNF: 08-MAY-95.
 DR MEROPS: M04.001; -.
 DR Interpro: IPR001570; Peptidase_M4.
 DR Interpro: IPR000130; Zn_MTPeptidse.
 DR Pfam: PF01447; Peptidase_M4; 1.
 DR Pfam: PF02868; Peptidase_M4_C; 1.
 DR PRINTS: PR00730; THERMOLYSIN.
 DR PROSITE: PS00142; ZINC_PROTEASE; 1.
 KW Hydrolyase; Metalloprotease; Zinc; Calcium; 3D-structure.
 FT METAL 142 142 ZINC (CATALYTIC).
 FT ACT_SITE 142 142 ZINC (CATALYTIC).
 FT METAL 146 146 ZINC (CATALYTIC).
 FT METAL 166 166 ZINC (CATALYTIC).
 FT ACT_SITE 231 231 PROTON DONOR.
 FT STRAND 4 11
 FT TURN 13 14
 FT STRAND 17 24
 FT STRAND 28 29
 FT STRAND 31 32
 FT TURN 36 37
 FT STRAND 39 43
 FT TURN 45 46
 FT STRAND 53 54
 FT STRAND 56 57
 FT STRAND 61 62
 FT HELIX 65 88
 FT TURN 92 94
 FT STRAND 100 106
 FT TURN 107 109
 FT STRAND 113 115
 FT STRAND 120 123
 FT STRAND 130 133
 FT HELIX 133 135
 FT HELIX 137 151
 FT TURN 152 153
 FT HELIX 159 179
 FT TURN 180 181
 FT STRAND 187 188
 FT TURN 190 192
 FT STRAND 193 193
 FT TURN 196 196
 FT TURN 198 199
 FT TURN 203 204
 FT HELIX 208 211
 FT TURN 212 212
 FT HELIX 217 219
 FT HELIX 225 229
 FT TURN 230 233

FT HELIX 234 246
 FT STRAND 248 250
 FT TURN 251 252
 FT STRAND 253 255
 FT HELIX 260 273
 FT TURN 274 274
 FT TURN 277 278
 FT HELIX 281 296
 FT TURN 298 299
 FT HELIX 301 312
 FT TURN 313 314
 SQ SEQUENCE 316 AA; 34333 MW; 004EDAF478744BB4 CRC64;
 Alignment Scores:
 Pred. No.: 1.3 Length: 316
 Score: 79.50 Matches: 40
 Percent Similarity: 41.98% Conservative: 28
 Best Local Similarity: 24.69% Mismatches: 53
 Query Match: 10.88% Indels: 41
 Ds: 1 Gaps: 10
 US-09-980-370-5 (1-402) x THER_BACSTH (1-316)
 QY 25 TCCTCTACTCAT-----GGAGAGCCAGCCGGAATGAGCT-----ACTATC 66
 DB 102 SerSerValHisTyrSerGlnGlyTyrAsnAsnAlaPheTyrPasnGlySerGluMetVal 121
 QY 67 GATGCCGATGGCAGTGGTGGCTGCTGCTGCTCCGACAAATGGACTGGCTC 126
 DB 122 TyrclyaspGlyaspGlyGlnThrPheIleProlseuSerGlyGly---Ileasp---Val 139
 QY 127 ACAGCTCATGGCGGCGCAACGAGTACGCTCTTCATGCAATGCAATGGCTTCAAT 186
 DB 140 ValAlaHisIsoleuThrHisAlaValThrAspTyrThr---AlaGlyLeuIleTyrGln 158
 QY 187 CCTGATACATATCTCATCTCAAG-----GATGTTACAGCGCAACGAGTAAAG 237
 DB 159 AsnIleuSerGlyAlaIleAsnGlnAlaIleSerAspIlePheGly---ThrIleValGlu 177
 QY 238 TACTACTATCCAGTCAACGACGAGCTTT-----CCCGGG--- 270
 DB 178 PheTyrAlaAsnIleAsnProAspTyrPheGlyGluAspValTyrThrProGlyIle 197
 QY 271 -----GATCACTAT 279
 DB 198 SerGlyAspSerLeuArgSerMetSerAspProAlaLysTyrGlyAspProAspHisTyr 217
 QY 280 GCGGTGATGATCTCCAGACAGCGGCGACGCGGAGACTTC-----ACGGTTGTTTC 333
 DB 218 SerLysArgTyrThrGlyThrGlnAspAsnGlyValHisIleAsnSerGlyIleIle 237
 QY 334 GAAGAAACGCGCTAACGGAATTAATAAGCGCGGAGCAGACATTCGGCTTTCACAGGAAGCC 393
 DB 238 AsnLysAlaAlaTyrIleuIleSerGlnGlyTyrHisTyrGlyValSerValIleGly 257
 QY 394 AATGGC 399
 DB 258 IleGly 259
 RESULT 5
 NPRS_BACST
 ID NPRS_BACST STANDARD; PRT: 551 AA.
 AC PA3133;
 DT 01-NOV-1995 (Rel. 32, Created).
 DT 01-NOV-1995 (Rel. 32, Last sequence update).
 DT 01-OCT-1996 (Rel. 34, Last annotation update).
 DE BacilloIysin precursor (EC 3.4.24.28) (Neutral protease).
 GN NPRS OR NPRM.
 OS Bacillus stearothermophilus.
 OC Bacteria: Firmicutes; Bacillales; Geobacillus.
 OX NCBI_TaxID=1422;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 236-239.

CC STRAIN-TELNE;
 RX MEDLINE-90368536; PubMed-2203733;
 RA Nishiyama Y., Imanaka T.;
 RT "Cloning and nucleotide sequence of the Bacillus stearothermophilus
 RT neutral protease gene and its transcriptional activator gene.";
 RL J. Bacteriol. 172:4861-4869(1990).
 [2]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 236-240.
 RC STRAIN-MK232;
 RX MEDLINE-89235633; PubMed-3149972;
 RA Kubo M., Imanaka T.;
 RT "Cloning and nucleotide sequence of the highly thermostable neutral
 RT protease gene from Bacillus stearothermophilus.";
 RL J. Gen. Microbiol. 134:1883-1892(1988).
 CC -1- FUNCTION: THERMOSTABLE EXTRACELLULAR ZINC METALLOPROTEASE
 CC -1- CATALYTIC ACTIVITY: Similar, but not identical, to that of
 CC thermolysin.
 CC -1- COFACTOR: BINDS 1 ZINC ION AND FOUR CALCIUM IONS (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M4.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M34237; AAA2625.1; -;
 DR EMBL: M21663; AAB02774.1; -;
 DR HSSP: P00800; 1TRL.
 DR MEROPS: M04.018; -;
 DR InterPro: IPR005075; pep_M4_propep.
 DR InterPro: IPR001570; Peptidase_M4.
 DR InterPro: IPR000130; Zn_MTPeptide.
 DR Pfam: PF01447; Peptidase_M4; 1.
 DR Pfam: PF02868; Peptidase_M4_C; 1.
 DR Pfam: PF03413; pep_M4_propep; 1.
 DR PRINTS: PR00730; THERMOLYSIN.
 DR PROSITE: PS00142; ZINC_PROTEASE; 1.
 KM Hydroxylase; Metalloprotease; Zinc; Calcium; zymogen; signal.
 FT SIGNAL. 1 31
 FT PROPEP 32 235
 FT CHAIN 236 551
 FT METAL 377 378
 FT ACT_SITE 378 378
 FT METAL 381 381
 FT ACT_SITE 401 401
 FT ACT_SITE 466 466
 FT CONFLICT 10 14
 FT CONFLICT 24 24
 FT CONFLICT 29 33
 FT CONFLICT 114 114
 FT CONFLICT 124 124
 FT CONFLICT 134 134
 FT CONFLICT 261 261
 FT CONFLICT 463 463
 SQ SEQUENCE 551 AA; 60616 MW; FCF4B2B5A7870129 CRC64;
 Alignment Scores:
 Pred. No.: 1.41 Length: 551
 Score: 79.50 Matches: 40
 Percent Similarity: 41.98 Conservative: 28
 Best Local Similarity: 24.698 Mismatches: 53
 Query Match: 10.888 Indels: 41
 DB: 1 Gaps: 10
 US-09-980-370-5 (1-402) x NPRS_BACST (1-551)
 QY 25 TCTTCACATGAT-----GGAGAGCGACAGCGAGATGAGACT-----ACTATC 66
 ||||| ||| ||| ||| |||

DB 337 SerSerValHisTyrSerGlnGlyTyrAsnAsnAlaPheTyrPasnGlySerGlnMetVal 356
 QY 67 GATGCCGATGCGATGATGAGGATGGCTCTGCTGCTTCGACATTTGGATGGCTC 126
 ||||| ||||| : : : : : ||||| || : : : : :
 DB 357 TyrGlyAspLysPrgLysPrgLysInThrPheIleProLeuSerGlyGly---IleAsp---Val 374
 QY 127 ACAGCTCATGGCGGCGACCAAGCTAGTACGCTCTTCTCATGATGATGATGCTTTCAT 166
 ||||| ||||| : : : : : ||||| ||||| : : : : :
 DB 375 ValAlaHisGlnLeuThrHisAlaValThrAspTyrThr---AlaGlyLeuIleTyrGln 393
 QY 187 CCGATACATATCTCATCTCAAG-----GATGTTACAGCGGCAAGCTAAAG 237
 : : : : : ||||| : : : : : ||||| ||||| : : : : :
 DB 394 AsnGluSerGlyAlaIleAsnGlnAlaIleSerAspIlePheGly---ThrLeuValGlu 412
 QY 238 TACACATACGATCATCAAGCGCTTTT-----CCGCGG--- 270
 : : : : : ||||| : : : : : ||||| ||||| : : : : :
 DB 413 PheTyrAlaAsnLysAsnProAspTyrGluIleGlyGluAspValTyrThrProGlyIle 432
 QY 271 -----GATCACTAT 279
 DB 433 SerGlyAspSerLeuArgSerMetSerAspProAlaTyrGlyAspProAspHisTyr 452
 QY 280 GCGGTGATGATCTCAAGACGCGGACGACGCGGAGACTTC-----ACGTTGTTTTC 333
 : : : : : ||||| : : : : : ||||| ||||| : : : : :
 DB 453 SerLysArgTyrThrGlyThrGlnAspAsnGlyValHisIleAsnSerGlyIleIle 472
 QY 334 GAGCAAGCGCTTACCGATTAATTAAGGCGGACGACGATTCGCTTCCACGGAAGCC 393
 : : : : : ||||| : : : : : ||||| ||||| : : : : :
 DB 473 AsnLysAlaValTyrLeuIleSerGlnGlyGlyThrHisTyrGlyValSerValValGly 492
 QY 394 AATGCC 399
 |||
 DB 493 IleGly 494
 RESULT 6
 GBLM_STRTR STANDARD; PRT; 348 AA.
 ID GBLM_STRTR
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Aldose 1-epimerase (EC 5.1.3.3) (Mutarotase).
 GN GBLM.
 OS Streptococcus thermophilus.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1308;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-A147;
 RX MEDLINE-90299833; PubMed-1694527;
 RX Poolman B., Royer T.J., Mainzer S.E., Schmidt B.F.;
 RT "Carbohydrate utilization in Streptococcus thermophilus:
 RT characterization of the genes for aldose 1-epimerase (mutarotase) and
 RT UDPglucose 4-epimerase.";
 RL J. Bacteriol. 172:4037-4047(1990).
 RL -1- FUNCTION: Mutarotase converts alpha-aldose to the beta-anomer. It
 CC is active on D-glucose, L-arabinose, D-xylose, D-galactose,
 CC maltose and lactose (by similarity).
 CC -1- CATALYTIC ACTIVITY: Alpha-D-glucose = beta-D-glucose.
 CC -1- PATHWAY: Hexose metabolism.
 CC -1- SIMILARITY: BELONGS TO THE ALDOSE EPIMERASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M38175; AAA26945.1; -;
 DR PIR: B44509; B44509.

RP MUTAGENESIS, AND DOMAINS.
 RX MEDLINE-91312880; PubMed-1677466;
 RA Py B., Bortoli-German I., Halech J., Chippaux M., Barras F.;
 RT "Cellulase ECG of *Erwinia chrysanthemi*: structural organization and
 RT importance of His98 and Glu13 residues for catalysis.";
 RL Protein Eng. 4:325-333(1991).
 RN [4]
 RP STEREOCHEMISTRY OF THE REACTION.
 RX MEDLINE-92225124; PubMed-1563515;
 RA Barras F., Bortoli-German I., Bauzen M., Rouvier J., Gey C.,
 RT Heyraud A., Hentisat B.;
 RT "Stereochemistry of the hydrolysis reaction catalyzed by
 RT endoglucanase Z from *Erwinia chrysanthemi*.";
 RL FEBS Lett. 300:145-148(1992).
 RN [5]
 RP STRUCTURE BY NMR OF 365-426.
 RX MEDLINE-98070232; PubMed-9405041;
 RA Brun E., Morland F., Gans P., Blackledge M.J., Barras F.,
 RT "Solution structure of the cellulose-binding domain of the
 RT endoglucanase Z secreted by *Erwinia chrysanthemi*.";
 RL Biochemistry 36:16074-16086(1997).
 RN [6]
 RP X-RAY CRYSTALLOGRAPHY (2.30 ANGSTROMS) OF 44-335.
 RX MEDLINE-21392910; PubMed-11501955;
 RA Chapon V., Czjzek M., El Hassouni M., Py B., Juy M., Barras F.;
 RT "Type II protein secretion in gram-negative pathogenic bacteria: the
 RT study of the structure/secretion relationships of the cellulase Cel5
 RT (formerly Ecg2) from *Erwinia chrysanthemi*.";
 RL J. Mol. Biol. 310:1055-1066(2001).
 CC -1- FUNCTION: REPRESENTS 97% OF THE GLOBAL CELLULOSE ACTIVITY.
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic
 CC linkages in cellulose.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: BELONGS TO CELLULOSE FAMILY A (FAMILY 5 OF GLYCOSYL
 CC HYDROLASES).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 DR EMBL: Y00540; CAAG604.1; -
 DR PIR: S03767; S03767.
 DR PDB: 1A1W; 06-MAY-98.
 DR PDB: 1EGZ; 26-MAR-99.
 DR InterPro: IPR003610; CBM_5_12.
 DR InterPro: IPR001547; GH_5_12.
 DR Pfam: PF00150; cellulase; 1.
 DR Pfam: PF02839; CBM_5_12; 1.
 DR SMART: SM00455; ChBD3; 1.
 DR PROSITE: PS00659; GLYCOSYL_HYDROL_F5; 1.
 KM Cellulose degradation; Hydrolyase; Glycosidase; Signal; 3D-structure.
 FT SIGNAL 1 43
 FT CHAIN 44 426
 FT DOMAIN 44 332
 FT DOMAIN 333 366
 FT DOMAIN 367 426
 FT ACT_SITE 176 176
 FT ACT_SITE 263 263
 FT DISULFID 368 425
 FT MUTAGEN 141 141
 FT MUTAGEN 176 176
 FT CONFLICT 293 295
 FT CONFLICT 350 364
 FT CONFLICT 388 426
 FT SEQUENCE 426 AA; 46418 MW; E78F2EE021FC5DA CRC64;

Alignment Scores:
 Pred. No.: 1.97 Length: 426
 Score: 78.00 Matches: 30
 Percent Similarity: 44.55% Conservative: 15
 Best Local Similarity: 29.70% Mismatches: 24
 Query Match: 10.67% Indels: 32
 DB: 1 Gaps: 7
 US-09-980-370-5 (1-402) x GUNZ_ERMCH (1-426)
 QY 7 TTCACGGAACGTCGAGCTCTTCTACTCATGAGAGGCA----- 45
 Db 232 TyrThrLeuHisPheTyrAlaGlyThrHisGlyGlnSerLeuArgAsnLysAlaArgGln 251
 QY 46 -----CCAGCGAATGACTACTATCATGATCGCATGCG 78
 Db 252 AlaLeuAsnAsnGlyIleAlaLeuPheValThrGluTrpGlyThrValAsnAlaAspGly 271
 QY 79 GATGGTGAAGGTGGCTGCTGCTGCTGCTGCGACATGAGAC---TGCTCACAGCTCAT 135
 Db 272 AsnGlyGly-----ValAsnGlnThrGluThrAspAlaTrpValThrPheMet 287
 QY 136 GCGCGACACCAAGCTGACCTGTTCTCATGAGATGCAATGCTTGAATCCTGATAC 195
 Db 288 ArgAspAsnAsn---IleSerAsnAlaAsnTrp-----AlaLeuAsnAspLysAsn 303
 QY 196 -----TATCTCATCTCAAGATGTTTACA---GGCGCAACGAGGTA 234
 Db 304 GluGlyAlaSerThrTyrTyrTrpAspSerLysAsnLeuThrGlnSerGlyLysVal 323
 QY 235 ANG 237
 Db 324 Lys 324
 RESULT 9
 ID NPRE_BACBR STANDARD; PRT: 527 AA.
 AC P43263;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE Bacillolysins precursor (EC 3.4.24.28) (Neutral protease).
 GN NFR.
 OS Bacillus brevis.
 OC Bacteria; Firmicutes; Bacillales; Paenibacillaceae; Brevibacillus.
 OX NCBI_Taxid-1393;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SRRAIN-7882;
 RX MEDLINE-91148616; PubMed-2290427;
 RA Anakov A.S., Bolotin A.P., Sorokin A.V.;
 RT "Structure of the Bacillus brevis metalloprotease gene.";
 RL Mol. Biol. (Mosk) 24:1363-1372(1990).
 RN [12]
 RP SEQUENCE OF 224-228, AND CHARACTERIZATION.
 RX MEDLINE-91148618; PubMed-2127074;
 RA Kaidalova N.V., Akinkina T.V., Khodova O.D., Kostrov S.V.,
 RT Strongin A.V.;
 RT "Analysis of the structure of Bacillus brevis neutral protease and
 RT its biosynthesis in Bacillus subtilis cells.";
 RL Mol. Biol. (Mosk) 24:1381-1392(1990)
 CC -1- FUNCTION: THERMO-LABILE EXTRACELLULAR ZINC METALLOPROTEASE.
 CC -1- CATALYTIC ACTIVITY: Similar, but not identical, to that of
 CC thermolysin.
 CC -1- CORRECTOR: BINDS 1 ZINC ION AND FOUR CALCIUM IONS (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M4.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

Db 84 ValaspaValThrPheSerGlyThrProAspGlyVal 95

RESULT 11

ITAV_CHICK STANDARD: PRT: 1034 AA.

AC P26008; 01-MAY-1992 (Rel. 22, Created)

DT 01-MAY-1992 (Rel. 22, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Integrin alpha-V precursor (Vitronectin receptor alpha subunit).

GN ITGAV.

OS Gallus gallus (Chicken).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; OC Gallus.

OX NCBI_TaxID=9031;

RN (1)

RP SEQUENCE FROM N.A.

RC TISSUE=Embryo;

RA MEDLINE=91104936; PubMed=1703004;

RA Bossy B., Reichardt L.F.;

RT "Chick integrin alpha V subunit molecular analysis reveals high conservation of structural domains and association with multiple beta subunits in embryo fibroblasts."

RL Biochemistry 29:10191-10198(1990).

CC -1- FUNCTION: THE ALPHA-V INTEGRINS ARE RECEPTORS FOR VITRONECTIN, CYTOTACTIN, FIBRONECTIN, FIBRINOGEN, LAMININ, MATRIX METALLOPROTEINASE-2, OSTEOPOINTIN, PROTHROMBIN, THROMBOSPONDIN AND VON WILLEBRAND FACTOR. THEY RECOGNIZE THE SEQUENCE R-G-D IN A WIDE ARRAY OF LIGANDS.

CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. THE ALPHA SUBUNIT IS COMPOSED OF AN HEAVY AND A LIGHT CHAIN LINKED BY A DISULFIDE BOND. ALPHA-V ASSOCIATES WITH EITHER BETA-1, BETA-3, BETA-5, BETA-6 OR BETA-8.

CC -1- SUBCELLULAR LOCATION: Type I membrane protein.

CC -1- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.

CC -1- SIMILARITY: CONTAINS 7 FG-GAP REPEATS.

CC -----

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

CC EMBL: M60517; AAA69138.1; -

CC PIR: A36108; A36108.

CC HSSP: P06756; 1JV2.

DR InterPro: IPR000413; Integrin_alpha.

DR Pfam: PF00357; Integrin_A; 1.

DR Pfam: PF01839; FG-GAP; 5.

DR PRINTS: PRO1185; INTEGRINA.

DR SMART: SM00191; InL_alpha; 5.

DR PROSITE: PS00242; INTEGRIN_ALPHA; 1.

DR Integrin: Cell adhesion; Receptor; Glycoprotein; Transmembrane; KW Signal; Repeat; Calcium.

FT SIGNAL 1 19 POTENTIAL.

FT CHAIN 20 1034 INTEGRIN ALPHA-V.

FT CHAIN 20 875 INTEGRIN ALPHA-V HEAVY CHAIN (POTENTIAL).

FT CHAIN 877 1034 INTEGRIN ALPHA-V LIGHT CHAIN (POTENTIAL).

FT DOMAIN 20 978 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 979 1002 POTENTIAL.

FT DOMAIN 1003 1034 CYTOPLASMIC (POTENTIAL).

FT REPEAT 35 96 FG-GAP 1.

FT REPEAT 101 166 FG-GAP 2.

FT REPEAT 167 224 FG-GAP 3.

FT REPEAT 237 290 FG-GAP 4.

FT REPEAT 291 350 FG-GAP 5.

FT REPEAT 355 415 FG-GAP 6.

FT REPEAT 420 473 FG-GAP 7.

FT CA_BIND 248 256 POTENTIAL.

FT CA_BIND 302 310 POTENTIAL.

FT CA_BIND 367 375 POTENTIAL.

FT CA_BIND 431 439 POTENTIAL.

FT SITE 1005 1009 GEFER MOTIF.

FT DISULFID 77 85 BY SIMILARITY.

FT DISULFID 126 146 BY SIMILARITY.

FT DISULFID 160 173 BY SIMILARITY.

FT DISULFID 479 488 BY SIMILARITY.

FT DISULFID 494 551 BY SIMILARITY.

FT DISULFID 612 618 BY SIMILARITY.

FT DISULFID 684 697 BY SIMILARITY.

FT DISULFID 838 900 INTERCHAIN (BY SIMILARITY).

FT DISULFID 895 900 BY SIMILARITY.

FT CARBOHYD 62 62 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 278 278 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 284 284 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 540 540 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 601 601 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 690 690 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 821 821 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 837 837 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 860 860 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 931 931 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 951 951 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 959 959 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 966 966 N-LINKED (GLCNAC. . .) (POTENTIAL).

SO SEQUENCE 1034 AA; 114388 MW; D76B08BA692DC684 CRC64;

Alignment Scores:

Pred. No.: 4.14 Length: 1034

Score: 75.50 Matches: 36

Percent Similarity: 37.75% Conservative: 21

Best local Similarity: 23.84% Mismatches: 47

Query Match: 10.33% Indels: 47

DB: 1 Gaps: 9

US-09-980-370-5 (1-402) x ITAV_CHICK (1-1034)

OY 19 TTGAGCTCTTCTACTCATGGAGAG---GCACGACGGAATGACTACTATGATCCGCAT 75

DB 149 PheAspValSerLysSerValGluTyrIleProCysArgSerThrThrIleAspAlaAsp 168

OY 76 GCGCATGTGAGGGGTGGCTGTCTCTCCGACAATGAGCTGCTCAGACATCAT 135

DB 169 GlyIleGly-----PheCysGlnGlyGlyPheSerIleAspPheThrLys----- 183

OY 136 GCGGCGACACAGTAGTA-----AGCTTTTTCATGGAATGGAATGGCTTTTG 183

DB 184 ---GlyAspArgValIleLeuGlyGlyProGlySerPheTyrTrpGlnGlyIleuIle 202

OY 184 AATGCT----- 189

DB 203 SerAspArgValAlaGluIleLeuAlaLysTyrAspSerLysValTyrSerThrLysTyr 222

OY 190 GATPACTATCTCATCTCAAGAGATGTTACAGGCGCACGAAGTAAGTACTACTATCA 249

DB 223 AspAspIleuAlaThrArgProIleSerAlaIaIa----- 234

OY 250 GTCAACGACGGTTTCCCGGGGATCACTATGCGGTGATGATCTTCAAGACGGCAGAAC 309

DB 235 PheAspAspSerPheTyrLeuGly-----TyrSerValAlaVal----- 246

OY 310 GCGGAGACCTTCAGG---GTTGTTTTCGAGGAAGACGCTTAACGGAATTAATAGCGCGCA 366

DB 247 ---GlyAspPheSerGlyAspGlyIleGluAspPheValSerGlyValProAlaGluAla 265

OY 367 GCAAGATTGCTCTTTCACGGAAGCAATGCG 399

DB 266 ArgThrLeuGlyMetValSerIleTyrAsnGly 276

RESULT 12

THR_BACST STANDARD: PRT: 548 AA.

DE Putative thiosulfate sulfurtransferase SSEA (EC 2.8.1.1).
GN SSEA OR RV3383 OR MT3382 OR MTCY1.23.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Fellwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RA "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=COC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umaya M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Blahai W.;
RA "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RL Submitted (Apr-2003) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: Thiosulfate + cyanide = sulfite + thiocyanate.
CC -1- SIMILARITY: BELONGS TO THE RHODANSE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: 292771; CAB07066.1;
CC DR EMBL: AE007147; AAK47725.1;
CC DR HSRP: P52197; 1E0C.
CC DR TIGR: MT3382;
CC DR Tuberculist: RV3283;
CC DR InterPro: IPR001763; Rhodanese-1like.
CC DR InterPro: IPR001307; Rhodanese.
CC DR Pfam: PF00581; Rhodanese; 2.
CC DR SMART: SM00450; RHOD. 2.
CC DR PROSITE: PS00380; RHODANSE_1; 1.
CC DR PROSITE: PS00683; RHODANSE_2; 1.
CC DR Hypothetical protein; transferase; Complete proteome.
CC KW ACT_SITE 245 245 BY SIMILARITY.
CC FT SITE 245 245
CC SQ SEQUENCE 297 AA: 33320 MW: 5930ABSP9C63A533 CRC64;

Alignment Scores:
Pred. No.: 5.7 Length: 297
Score: 73.50 Matches: 22
Percent Similarity: 45.35% Conservative: 17
Best Local Similarity: 25.58% Mismatches: 30
Query Match: 10.05% Indels: 17
DB: 1 Gaps: 4

US-09-980-370-5 (1-402) x THF2.MYCTU (1-297)
QY 118 GACTGCTCAGACGCTATGCGGACCAACAGTAGAAGCTTTCTCATGGAATGGAATG 177
DB 24 Asprtrplensealshmetglyala-----Proglyleu 35
QY 178 GCCTTG---AATCTCATATCATCTCATCTCAAGAGATGTT-----ACAGGCGCA 225

DB 36 AlaIleValGluSerIaspGluAspValLeuLeuTyrAspValGlyHisIleProGlyAla 55
QY 226 ACGAAGCTAAAGTACTACTATCCAGTCAACGACGCT-----TTTCCCGGG 270
DB 56 ValIysIleAspTrpHisThrAspLeuAsnAspProArgValArgAspTyrIleAsnGly 75
QY 271 GATCAGTATCGCGGTGATGATCTCCCAAGACGGGACGAACCGCGGAGACTTCACGGTGT 330
DB 76 GluGlnPheAlaGluLeuMetAspArgLysGlyIleAlaArgAspAspThrValValIle 95
QY 331 TTCGAAGAAACCGCTTAC 348
DB 96 TyrGlyAspLysSerAsn 101

Search completed: June 26, 2003, 00:44:53
Job time : 20 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus.n2p model

Run on: June 25, 2003, 23:06:58 : Search time 35.5 seconds
(without alignments)
4666.532 Million cell updates/sec

Title: US-09-980-370-5

Perfect score: 731
Sequence: 1 gcaagcttcacggaagcgtt.....ccacggaagcgaagcgcgc 402

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 1343160

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+.n2p.model -DEV=xlh
-O=/gen2.1/USPTO.spool/US09980370/runat.19062003.174442.8856/app_query.fasta.1.583
-DB=SPRMBL_21 -QFMT=fastan -SUFFIX=trspc -MINMATCH=0.1 -LOOPEC=0 -LOOEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human4.0.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pic -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09980370.eccn.1.1.57.grunat.19062003.174442.8856 -NCPu=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

SPRMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_minc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	DB	ID	Description
1	716	97.9	1223	2	Q9ZNB5	Q9ZNB5 porphyromon

2	716	97.9	1687	2	Q9R9B7	Q9R9B7 porphyromon
3	716	97.9	1704	2	Q51816	Q51816 porphyromon
4	715	97.8	1097	2	P72196	P72196 porphyromon
5	715	97.8	1732	2	O52050	O52050 porphyromon
6	715	97.8	1732	2	O51817	O51817 porphyromon
7	715	97.8	1732	2	O07442	O07442 porphyromon
8	714	97.7	1706	2	O51839	O51839 porphyromon
9	711	97.3	1723	2	P72197	P72197 porphyromon
10	709	97.0	1358	2	P96967	P96967 porphyromon
11	709	97.0	1723	2	P72194	P72194 porphyromon
12	706	96.6	1706	2	O51838	O51838 porphyromon
13	705	96.0	925	2	O9F4J0	O9F4J0 porphyromon
14	180	24.6	312	2	O9K1B3	O9K1B3 porphyromon
15	90.5	12.4	218	16	O9J215	O9J215 neisseria m
16	90.5	12.4	286	17	O9V211	O9V211 pyrococcus
17	89	12.2	418	9	O80060	O80060 staphylococ
18	89	12.2	1225	9	O8SDK2	O8SDK2 staphylococ
19	89	12.2	1261	9	O9MBN9	O9MBN9 staphylococ
20	88.5	12.1	218	16	O9J0U0	O9J0U0 neisseria m
21	88.5	12.1	1217	16	O8Y378	O8Y378 raietonia s
22	87	11.9	910	3	P87211	P87211 orpinomyces
23	85.5	11.7	2275	17	O8TSE8	O8TSE8 methanocarc
24	84	11.5	494	2	O86099	O86099 alteromonas
25	83.5	11.4	256	17	O8TPD2	O8TPD2 methanocarc
26	83.5	11.4	548	2	O45779	O45779 bacillus th
27	83.5	11.4	603	10	O9M653	O9M653 polyonatum
28	83	11.4	373	16	O3926	O3926 mycobacteri
29	83	11.4	806	3	O9P478	O9P478 agaricus bl
30	82.5	11.3	536	17	O97YP2	O97YP2 sulfolobus
31	81.5	11.1	366	17	O26770	O26770 methanobact
32	81	11.5	355	10	O9AUL4	O9AUL4 oryza sativ
33	81	11.1	1372	2	O54151	O54151 shigella fl
34	81	11.1	1373	2	O9A158	O9A158 shigella fl
35	79.5	10.9	534	5	O25402	O25402 lymanaea sta
36	79.5	10.9	735	16	O8XNP9	O8XNP9 clostridium
37	78.5	10.7	348	2	O8VS90	O8VS90 streptococ
38	78.5	10.7	448	2	O8VS84	O8VS84 streptococ
39	78	10.7	499	2	O45532	O45532 bacillus su
40	78	10.7	785	3	O9PBH5	O9PBH5 humicola in
41	78	10.7	803	10	O8W0T5	O8W0T5 sorghum bic
42	78	10.7	1606	12	O911D6	O911D6 white spot
43	77.5	10.6	438	3	O9P8K8	O9P8K8 leptosphaer
44	77.5	10.6	1039	5	O9VQB1	O9VQB1 drosophila
45	77.5	11.0	3014	12	O39928	O39928 hepatitis c

ALIGNMENTS

RESULT 1
Q9ZNB5 PRELIMINARY: PRT: 1223 AA.
AC Q9ZNB5: 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE 130K-HMGD (Fragment)
OS Porphyromonas gingivalis (Bacteroides gingivalis)
OC Bacteria; CFB group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
OC Porphyromonas.
OX NCBI_TaxID=837;
RP [1]
RP SEQUENCE FROM N.A.
RC STRAIN=381;
RX MEDLINE=99143166; PubMed=9988746;
RA Shibata Y., Hayakawa M., Takiguchi H., Shiroza T., Abiko Y.:
RT "Determination and characterization of the hemagglutinin-associated
RT short motifs found in Porphyromonas gingivalis multiple gene
RT products."
RL J. Biol. Chem. 274:5012-5020(1999).
DR EMBL: AB019363; BAA34341.1; -
DR InterPro: IPR000977; DNA ligase.
DR InterPro: IPR001769; Peptidase_C25.
DR Pfam: PF01364; Peptidase_C25; 3.

DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
FT NON_TER 1 1
SQ SEQUENCE 1223 AA: 131542 MW: 00225CD2BA9F91B3 CRC64:

Alignment Scores:

Pred. No.:	2,48e-62	Length:	1223
Score:	716.00	Matches:	131
Percent Similarity:	99.25%	Conservative:	2
Best Local Similarity:	97.76%	Mismatches:	1
Query Match:	97.95%	Indels:	0
DB:	2	Gaps:	0

US-09-980-370-5 (1-402) x Q92NB5 (1-1223)

QY 1 GCAGACTTCACGGAACGTTGAGTCTTCTACTCATGTGAGAGCAGCAGCGGAATGACT 60
DB 648 AlaAspIleThrGluThrPheGluSerThrHisGluAlaProAlaGluTrpThr 667
QY 61 ACTATCGATGCCGATGGCGATGGTGGCTCTGCTCTCTTCGCGACAAATTGGAC 120
DB 668 ThrIleAspAlaAspGlyAspGlyGlnGlyTrpLeuCySleuSerSerGlyGlnLeuAsp 687
QY 121 TGGCTCAGCTCATGGCGGCGGACCAACGTAAGCTTCTTCATGATGATGATGCT 180
DB 688 TrpLeuThrAlaHisGlyGlyThrAsnValValAlaSerPheSerTrpAsnGlyMetAla 707
QY 181 TTGAATCCTGATTAATCTATCTCATCTCAAGAGTTTACAGCGGCAAGAGTAAGTAC 240
DB 708 LeuAsnProAspAsnTyrLeuIleSerLysAspValThrGlyAlaThrLysValLysTyr 727
QY 241 TACTATCCAGTCAACGAGCGTTTCCCGGATCACTATGCGGTGATGATCTCCAGAGC 300
DB 728 TyrTyrAlaValAlaAsnAspGlyPheProGlyAspHisTyrAlaValMetIleSerLysThr 747
QY 301 GGCACGAACGCCGAGACTTCAAGGTTTTCGAGAAGAACGCCCTAAGCAATTAATAG 360
DB 748 GlyThrAsnAlaGlyAspPheThrValValPheGluGluThrProAsnGlyIleAsnLys 767

OY 361 GCGGAGCAGCAAGTTGGCTTTTCACAGGAAAGCCAAATGGCGCC 402

DB 768 GlyGlyAlaArgPheGlyLeuSerThrGluAlaAsnGlyAla 781

RESULT 2

Q9R9B7 PRELIMINARY: PRT: 1687 AA.

AC 09R9B7: AC 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hemagglutinin/protease.
CN HAGE.
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; CFb group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
OC Porphyromonas.
OX NCBL_TaxID=837;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=381.
RA Han N., Dong H., Proguiske-Fox A.;
RT "Cloning and characterization of hage from P. gingivalis 381."
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF026946; AAD01810.1; -
DR HSSP: P95493; 1CVR.
DR MEROPS: C25.001; -
DR InterPro: IPR000977; DNA_Ligase.
DR InterPro: IPR001769; Peptidase_C25.
DR Pfam: PF01364; Peptidase_C25; 3.
DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
KW PROTEINASE.
SQ SEQUENCE 1687 AA: 183702 MW: D085B516A39FE70 CRC64:

Alignment Scores: 2.61e-62 Length: 1687
Pred. No.: 2.61e-62 Length: 1687

Score: 716.00 Matches: 131
Percent Similarity: 99.25% Conservative: 2
Best Local Similarity: 97.76% Mismatches: 1
Query Match: 97.95% Indels: 0
DB: 2 Gaps: 0

US-09-980-370-5 (1-402) x Q9R9B7 (1-1687)

QY 1 GCAGACTTCACGGAACGTTGAGTCTTCTACTCATGTGAGAGCAGCAGCGGAATGACT 60
DB 1140 ThrIleAspAlaAspGlyAspGlyGlnGlyTrpLeuCySleuSerSerGlyGlnLeuAsp 1159
QY 121 TGGCTCAGCTCATGGCGGCGGACCAACGTAAGCTTCTTCATGATGATGATGCT 180
DB 1160 TrpLeuThrAlaHisGlyGlyThrAsnValValAlaSerPheSerTrpAsnGlyMetAla 1179
QY 181 TTGAATCCTGATTAATCTATCTCATCTCAAGAGTTTACAGCGGCAAGAGTAAGTAC 240
DB 1180 LeuAsnProAspAsnTyrLeuIleSerLysAspValThrGlyAlaThrLysValLysTyr 1199
QY 241 TACTATCCAGTCAACGAGCGTTTCCCGGATCACTATGCGGTGATGATCTCCAGAGC 300
DB 1200 TyrTyrAlaValAlaAsnAspGlyPheProGlyAspHisTyrAlaValMetIleSerLysThr 1219
QY 301 GGCACGAACGCCGAGACTTCAAGGTTTTCGAGAAGAACGCCCTAAGCAATTAATAG 360
DB 1220 GlyThrAsnAlaGlyAspPheThrValValPheGluGluThrProAsnGlyIleAsnLys 1239
QY 361 GCGGAGCAGCAAGTTGGCTTTTCACAGGAAAGCCAAATGGCGCC 402
DB 1240 GlyGlyAlaArgPheGlyLeuSerThrGluAlaAsnGlyAla 1253

RESULT 3

Q51816 PRELIMINARY: PRT: 1704 AA.

ID Q51816: ID 01-NOV-1996 (TREMBLrel. 01, Created)
AC Q51816: AC 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Arg-gingipain-1 proteinase.
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; CFb group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
OC Porphyromonas.
OX NCBL_TaxID=837;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95138080; PubMed=7836351;
RA Pavloff N., Potempa J., Pike R.N., Prochazka V., Kiefer M.C.,
RA Travis J., Barr P.J.;
RT "Molecular cloning and structural characterization of the Arg-gingipain proteinase of Porphyromonas gingivalis. Biosynthesis as a
RT proteinase-adesin polypeptide."
RL J. Biol. Chem. 270:1007-1010(1995).
DR EMBL: U15282; AAA69539.1; -
DR HSSP: P95493; 1CVR.
DR MEROPS: C25.001; -
DR InterPro: IPR000977; DNA_Ligase.
DR InterPro: IPR001769; Peptidase_C25.
DR Pfam: PF01364; Peptidase_C25; 3.
DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
FT CHAIN 228 719 MATURE 50-KDA CYSTEINE PROTEINASE
FT GINGIPAIN.
SQ SEQUENCE 1704 AA: 185436 MW: 6A34B0131C2A676 CRC64:

Alignment Scores: 2.62e-62 Length: 1704
Score: 716.00 Matches: 131
Percent Similarity: 99.25% Conservative: 2
Best Local Similarity: 97.76% Mismatches: 1

Query Match: 97.95% Idents: 0
 DB: 2 Gaps: 0
 US-09-980-370-5 (1-402) x 051816 (1-1704)

QY 1 GCAGACTTCACGGAAACGTTTCAGTCTTCTACTCATGAGAGACACAGCGAATGACT 60
 DB 1137 ATASpPheHrGluThrPheGluSerThrHisGlyAlaProAlaGluTrpPhe 1156

QY 61 ACTATCGATCGCATGCGATGCGATGCGATGCGATGCGATGCGATGCGATGCGATGCGAT 120
 DB 1157 ThrLeuPalaAspGlyAspGlyLeuGlyLeuGlyLeuGlyLeuGlyLeuGlyLeuGly 1176

QY 121 TGGCTCAGCTCATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
 DB 1177 TrpLeuThrAlaHisGlyGlyThrAlaValAlaSerPheSerTrpAsnGlyMetAla 1196

QY 181 TTGAATCCGATTAATCTCATCTCAAGGATGTTACAGCGGCGGCGGCGGCGGCGGCGGCGG 240
 DB 1197 LeuAspProAspAsnTyLeuLeuSerLysAspValThrGlyAlaThrLysValLysTrp 1216

QY 241 TACTATCCAGTCAACGAGGTTTCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
 DB 1217 TyrTyAlaValAlaAspGlyPheProGlyAspHisTyAlaValMetLieserLysTrp 1236

QY 301 GGCAGGAGCGGCGGAGACTTCACGGTGTTCGAGAAAGCGGCGGCGGCGGCGGCGGCGGCGG 360
 DB 1237 GlyThrAsnAlaGlyAspPheThrValAlaPheGluGluThrProAsnGlyLysAsnLys 1256

QY 361 GCGCGGAGGATTCGCTTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 402
 DB 1257 GlyGlyAlaAspPheGlyLeuSerThrGluAlaAsnGlyAla 1270

RESULT 4
 ID P72196 PRELIMINARY: PRT: 1097 AA.
 AC P72196: 01-FEB-1997 (TREMBLrel. 02, Created)
 DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
 DR 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE TOMB-linked adhesin precursor.
 OS Porphyromonas gingivalis (Bacteroides gingivalis).
 OC Bacteria; CFB group; Bacteroidetes; Bacteroidales; Porphyromonadaceae.
 OC Porphyromonas.
 OX NCBI_TaxID=837;
 RN [1]
 RM SEQUENCE FROM N.A. PubMed=9244265;
 RA Aduse-Opoku J., Stanley J.M., Young K.A., Muir J., Rangarajan M.,
 RT "The tla gene of Porphyromonas gingivalis W50: a homologue of the
 RT arginine-specific protease precursor (PRP1) which shares sequence
 RT similarity to TOMB-linked receptors."
 RL J.Bacteriol. 179:4729-4738(1997).
 DR InterPro: IPR000977; DNA_Ligase.
 DR InterPro: IPR001769; Peptidase_C25.
 DR Pfam: PF01364; Peptidase_C25.
 DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
 KW Signal.
 ET SIGNAL.
 SO SEQUENCE 1097 AA; 118731 MW; 73BBA37BA21F8B9 CRC64;

Alignment Scores:
 Pred. No.: 3.07e-62 Length: 1097
 Score: 715.00 Matches: 131
 Percent Similarity: 99.25% Conservative: 2
 Best Local Similarity: 97.76% Mismatches: 1
 Query Match: 97.81% Gaps: 0
 DB: 2 Gaps: 0
 US-09-980-370-5 (1-402) x P72196 (1-1097)

QY 1 GCAGACTTCACGGAAACGTTTCAGTCTTCTACTCATGAGAGACACAGCGAATGACT 60
 DB 522 ATASpPheHrGluThrPheGluSerThrHisGlyAlaProAlaGluTrpPhe 541

QY 61 ACTATCGATCGCATGCGATGCGATGCGATGCGATGCGATGCGATGCGATGCGATGCGAT 120
 DB 542 ThrLeuPalaAspGlyAspGlyLeuGlyLeuGlyLeuGlyLeuGlyLeuGlyLeuGly 561

QY 121 TGGCTCAGCTCATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
 DB 562 TrpLeuThrAlaHisGlyGlyThrAlaValAlaSerPheSerTrpAsnGlyMetAla 581

QY 181 TTGAATCCGATTAATCTCATCTCAAGGATGTTACAGCGGCGGCGGCGGCGGCGGCGGCGG 240
 DB 582 LeuAspProAspAsnTyLeuLeuSerLysAspValThrGlyAlaThrLysValLysTrp 601

QY 241 TACTATCCAGTCAACGAGGTTTCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
 DB 602 TyrTyAlaValAlaAspGlyPheProGlyAspHisTyAlaValMetLieserLysTrp 621

QY 301 GGCAGGAGCGGCGGAGACTTCACGGTGTTCGAGAAAGCGGCGGCGGCGGCGGCGGCGGCGG 360
 DB 622 GlyThrAsnAlaGlyAspPheThrValAlaPheGluGluThrProAsnGlyLysAsnLys 641

QY 361 GCGCGGAGGATTCGCTTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 402
 DB 642 GlyGlyAlaAspPheGlyLeuSerThrGluAlaAsnGlyAla 655

RESULT 5
 ID 052050 PRELIMINARY: PRT: 1732 AA.
 AC 052050: 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DR 01-JUN-1997 (TREMBLrel. 12, Last annotation update)
 DE Lysine specific cysteine protease
 OS Porphyromonas gingivalis (Bacteroides gingivalis).
 OC Bacteria; CFB group; Bacteroidetes; Bacteroidales; Porphyromonadaceae.
 OC Porphyromonas.
 OX NCBI_TaxID=837;
 RN [1]
 RM SEQUENCE FROM N.A. PubMed=9632563;
 RA Lewis J.P., Macrina F.L.;
 RT "IS195, an insertion sequence-like element associated with protease
 RT genes in Porphyromonas gingivalis."
 RL Infect. Immun. 65:1052-1062(1998).
 DR InterPro: IPR007653; MEROPS: C25.002.
 DR InterPro: IPR000977; DNA_Ligase.
 DR InterPro: IPR001769; Peptidase_C25.
 DR Pfam: PF01364; Peptidase_C25.
 DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
 KW Peptidase.
 ET SIGNAL.
 SO SEQUENCE 1732 AA; 187931 MW; B2337463DCB5EAS CRC64;

Alignment Scores:
 Pred. No.: 3.3e-62 Length: 1732
 Score: 715.00 Matches: 131
 Percent Similarity: 99.25% Conservative: 2
 Best Local Similarity: 97.76% Mismatches: 1
 Query Match: 97.81% Gaps: 0
 DB: 2 Gaps: 0
 US-09-980-370-5 (1-402) x 052050 (1-1732)

QY 1 GCAGACTTCACGGAAACGTTTCAGTCTTCTACTCATGAGAGACACAGCGAATGACT 60
 DB 1157 ATASpPheHrGluThrPheGluSerThrHisGlyAlaProAlaGluTrpPhe 1176

QY 61 ACTATCGATCGCATGCGATGCGATGCGATGCGATGCGATGCGATGCGATGCGATGCGAT 120

QY	DB	1177	ThrlleasapalaspjlyaspjolynglytrpLeuCYleuSerSerCglyleuAsp	1196
QY	121	TGCGTACAGCTCATGCGCGGCAACGATACGATCTTCCTTCGATCGATCGCC	180	
DB	1197	TrpLeuTrlrlAlahSlgLYglYserAsnValValSerSerPheSerTrpAlmglyMetAla	1216	
QY	181	TTGATACCTGATTAACATCATCATCAATCAACAGCATTTACAGCGCGCAACAGATTAAGTAC	240	
DB	1217	LeuAsnTrpAspAsnTrpAlmglylIleSerSerLYaspaTrlthCgATlthValValLYsTrY	12356	
QY	241	TACATCAATCAAGCAAGAGGGTTTCGCGGGATACATCGATATGATATCTTCGATCAAGAC	300	
DB	1237	TYtrTAlaValAlaValAsnAspGlylPheTrpCgLYasplnslTylAlaValIleCylSerLYsTh	12556	
QY	301	GGCAGCAACGGCGGACACTTCACAGGCTTCCTTCGACAAACGGCTTACGATATTAATAC	360	
DB	1257	GlyThrAsnAlaAlaCgLYaspPheTrpValValPheCgLYlthCgTrpAlmglylIleAsnLYs	12766	
QY	361	GGCGAGCAACATATCGCTGCTTCCTTCACAGCAAGCAATGCGGCC	402	
DB	1277	GLYAlaAlaPheCgLYleuSerTrhCgLYAlaAsnLYsAla	1290	
RESULT 6				
ID	Q51817	PRELIMINARY:	PRF: 1732 AA.	
AC	O51817	1996 (TEMBLRef. 01, Created)		
DC	O1-NOV-1996	(TEMBLRef. 01, Sequence update)		
DT	01-JUN-2002	(TEMBLRef. 21, Last annotation update)		
DE	Porphyralin.			
CN	PRP.			
OS	Porphyromonas gingivalis (Bacteroides gingivalis).			
OC	Bacteria: CFB group: Bacteroidales: Bacteroidales: Porphyromonadaceae:			
OX	NCBI TaxID=837.			
RN	111	SEQUENCE FROM N.A.		
RP	STRAIN=12.			
RX	MEDLINE=66213011: PubMed=8631659;			
RT	Analysis of the prtP gene encoding porphyralin, a cysteine proteolase			
RA	proteolase from strain M-8 N., Patel J.M., Whitlock J.,			
RL	J. Bacteriol. 178:2734-2741(1996).			
EMBL	U42210: AAB05651.1; -			
DR	Interp: C35:PF000977: DNA, 1199se.			
DR	Interp: IPH001769: Peptidase_C25.			
DR	Pfam: PF01364: Peptidase_C25; 3.			
DR	PROSITE: PS00697: DNA_LIGASE_A1: UNKNOWN.1.			
SO	SEQUENCE 1732 AA: 187875 MW: 654271DBE67BCAE4 CRC64:			
Alignment Scores:				
	Prod. No:	3.3e-62	Length: 1732.	
	Percent Similarity:	715.00	Matches: 131.	
	Best Local Similarity:	99.25%	Conservative: 2	
	Query Match:	97.81%	Mismatches: 1	
	DB:	2	Indels: 0	
			Gaps: 0	
US-09-980-370-5 (1-403) x 051817 (1-1732)				
QY	1	GCAGACTTCGAGCAACCTTCGAGCTCTACTCATGAGAGAGACGACGAGATGACT	60	
DB	1157	AlaAspPheTrlrlTrlthPheGlySerTrhTrlslgLYAlaTrlAlaGlyTrpTh	11766	
QY	61	ACATGATCGATCGATCGATCGATCGATCGATCGATCGATCGATCGATCGATCGATCGATCGAC	120	
DB	1177	ThrlleasapalaspjlyaspjolynglytrpLeuCYleuSerSerCglyleuAsp	1196	
QY	121	TGCGTACAGCTCATGCGCGGCAACGATACGATCTTCCTTCGATCGATCGCC	180	

QY	1197	TRP	PHEN	THR	ALA	ISG	LYG	YSR	VAL	VAL	SESR	PHSE	STR	PAN	GLY	ME	1216
DB	1197	TRP	PHEN <td>THR</td> <td>ALA<td>ISG<td>LYG<td>YSR<td>VAL<td>VAL<td>SESR<td>PHSE<td>STR<td>PAN<td>GLY<td>ME<td>1216</td></td></td></td></td></td></td></td></td></td></td></td></td>	THR	ALA <td>ISG<td>LYG<td>YSR<td>VAL<td>VAL<td>SESR<td>PHSE<td>STR<td>PAN<td>GLY<td>ME<td>1216</td></td></td></td></td></td></td></td></td></td></td></td>	ISG <td>LYG<td>YSR<td>VAL<td>VAL<td>SESR<td>PHSE<td>STR<td>PAN<td>GLY<td>ME<td>1216</td></td></td></td></td></td></td></td></td></td></td>	LYG <td>YSR<td>VAL<td>VAL<td>SESR<td>PHSE<td>STR<td>PAN<td>GLY<td>ME<td>1216</td></td></td></td></td></td></td></td></td></td>	YSR <td>VAL<td>VAL<td>SESR<td>PHSE<td>STR<td>PAN<td>GLY<td>ME<td>1216</td></td></td></td></td></td></td></td></td>	VAL <td>VAL<td>SESR<td>PHSE<td>STR<td>PAN<td>GLY<td>ME<td>1216</td></td></td></td></td></td></td></td>	VAL <td>SESR<td>PHSE<td>STR<td>PAN<td>GLY<td>ME<td>1216</td></td></td></td></td></td></td>	SESR <td>PHSE<td>STR<td>PAN<td>GLY<td>ME<td>1216</td></td></td></td></td></td>	PHSE <td>STR<td>PAN<td>GLY<td>ME<td>1216</td></td></td></td></td>	STR <td>PAN<td>GLY<td>ME<td>1216</td></td></td></td>	PAN <td>GLY<td>ME<td>1216</td></td></td>	GLY <td>ME<td>1216</td></td>	ME <td>1216</td>	1216
QY	181	TTG	AAT	CTC	CGA	TAC	TAC	ATC	TAT	CTC	ATC	TCA	AA	GAG	TTC	TAC	240
DB	1217	LEU	ASN	PRO	ASP	ASN	TRP	LEU	ILE	SER	LYS	ASP	VAL	TRP	GLY	VAL	1236
QY	241	TAC	ATC	ATC	CGA	CAAC	GAG	GGT	TTT	CCG	GAG	TAC	ATC	ATC	CGG	TAT	300
DB	1237	Tyr	Tyr	Ala	Val	Ala	Asn	Asp	Gly	Phe	Pro	Gly	Asp	His	Trp	Ala	1256
QY	301	GGC	AGC	ACG	CGG	AGC	ATT	CGC	GGT	TTT	CTT	CGA	CAAC	GCG	TTC	AA	360
DB	1257	Gly	His	Asn	Ala	Asp	Gly	Asp	Pro	Thr	Val	Ala	Pro	Gly	Leu	Trp	1276
QY	361	GGG	GGG	AGC	AAC	ATC	TGC	GTG	CTT	TCC	CGG	AGC	CAAC	GAG	CGG	CC	402
DB	1277	GLY	GLY	ALA	TRP	PHEN	LYS	LEU	SER	THR	GLY	VAL	ALA	ASN	GLY	ALA	1290
RESULT 7																	
ID	007442	PRELIMINARY: PRT: 1732 AA.															
NC	007442	PRT: 1732 AA.															
NC	007442	PRT: 1732 AA.															
DT	01-JUN-1997	(TEMPREL: 04, Last sequence update)															
DT	01-JUN-1997	(TEMPREL: 04, Last sequence update)															
DE	01-JUN-2002	(TEMPREL: 21, Last annotation update)															
DT	Lysine-specific cysteine proteinase.																
GN	PRK.																
OS	Porphyromonas gingivalis (Bacteroides gingivalis).																
OC	Bacteroides																
OC	Bacteroides																
OC	Bacteroides																
OC	Bacteroides																
OC	Bacteroides																
OC	Bacteroides																
OC	Bacteroides																
OC	Bacteroides																

OY 241 TACTATCCAGTCAGACGCGTTTCCGGGATCACTATCGGTGATGATTCGAAGAC 300
 DB 1237 TTTTAAAlaValaAsnAspGlyPheProGlyAspHisTyrAlaValMetIleSerLysThr 1256
 OY 301 GGCAGCAGCGCCGACGACTTCACGGTGTTCGAGAAGACCTTAACGGAATAAATGAG 360
 DB 1257 GYTThraAsnAlaGlyAspPheThrValAlaPheGluGluThrProAsnGlyIleAsnLys 1276
 OY 361 GGGGAGCAAGATTCGCTTTCACAGGAAGCAATGGCGCC 402
 DB 1277 GYgGlyAlaArgPheGlyLeuSerThrGluAlaAsnGlyAla 1290

RESULT 8

O51839 PRELIMINARY; PRT; 1706 AA.
 AC O51839; O51840:
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Arginine-specific thiol protease precursor.
 GN PRT.
 OS Porphyromonas gingivalis (Bacteroides gingivalis).
 OC Bacteria; CFb group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
 OC Porphyromonas.
 OX NCBI_Taxid=837;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-W50;
 RX MEDLINE=95160709; PubMed=7857299;
 RA Kirsbaum L., Sotliropoulos C., Jackson C., Cleal S., Stakeski N.,
 RA Reynolds E.C.;
 RT "Complete nucleotide sequence of a gene prtr of Porphyromonas
 RT gingivalis M50 encoding a 132 kDa protein that contains an arginine-
 RT specific thiol endopeptidase domain and a haemagglutinin domain.";
 RL Biochem. Biophys. Res. Commun. 207:424-431(1995).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN-W50;
 RX MEDLINE=96311339; PubMed=8713096;
 RA Stakeski N., Cleal S.M., Reynolds E.C.;
 RT "Characterization of a Porphyromonas gingivalis gene prtr that encodes
 RT an arginine-specific thiol proteinase and multiple adhesins.";
 RL Biochem. Biophys. Res. Commun. 224:603-610(1996).
 RN (3)
 RP SEQUENCE FROM N.A.
 RC STRAIN-W50;
 RA Reynolds E.;
 RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.
 RN (4)
 RP SEQUENCE FROM N.A.
 RC STRAIN-W50;
 RA Stakeski N.;
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL: L26341; AAC18876.1; -.
 DR HSSP: P95493; ICVR.
 DR InterPro: IPR000977; DNA_Ligase.
 DR InterPro: IPR001769; Peptidase_C25.
 DR Pfam: PF01364; Peptidase_C25.3.
 DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
 KW Protease; Signal.
 FT SIGNAL 1 23 POTENTIAL.
 FT CHAIN 228 1706 ARGININE-SPECIFIC THIOLE PROTEASE.
 SO SEQUENCE 1706 AA; 185626 MW; E8BD07C981B844 CRC64;

Alignment Scores:

Pred. No.: 4.13e-62 Length: 1706
 Score: 714.00 Matches: 131
 Percent Similarity: 99.25% Conservative: 2
 Best Local Similarity: 97.76% Mismatches: 1
 Query Match: 97.67% Indels: 0
 DB: 2 Gaps: 0

US-09-980-370-5 (1-402) x O51839 (1-1706)

OY 1 GCAGACTTCACGGAACGCTTCGAGTCTTCTACTATGAGAGGACCGCAATGGACT 60
 DB 1139 AlaAspPheThrGluThrPheGluSerSerThrHisGlyGluAlaProAlaGluThrPthr 1158
 OY 61 ACTATGATGCGGATGGGATGGAGGTTGGCTCTCTCTCTCTCGCAATGGAC 120
 DB 1159 ThrIleAspAlaAspGlyAspGlyGlnGlyIlePheLysLeuSerSerGlyGlnLeuAsp 1178
 OY 121 TGGCTCAGACTCATGGCGGACCACTAGTAACCTTTCTCTCATGAAATGAAATGGCT 180
 DB 1179 TrpLeuThrAlaHisGlyGlyThrAsnValAlaSerSerPheSerTrpAsnGlyMetAla 1198
 OY 181 TTGAATCCGATTAACATCTATCTCATCAAGATGTTACAGCGGACCAAGAGTAAGTAC 240
 DB 1199 LeuAsnProAspAspTyrIleuIleSerLysAspAlaThrGlyAlaThrLysValLysTyr 1218
 OY 241 TACTATCCAGTCAGACGCGTTTCCGGGATCACTACTGCGGATGATTCACAGACG 300
 DB 1219 TTTTAAAlaValaAsnAspGlyPheProGlyAspHisTyrAlaValMetIleSerLysThr 1238
 OY 301 GGCAGCAGCGCCGACGACTTCACGGTGTTCGAGAAGACCTTAACGGAATAAATGAG 360
 DB 1239 GYTThraAsnAlaGlyAspPheThrValAlaPheGluGluThrProAsnGlyIleAsnLys 1258
 OY 361 GGGGAGCAAGATTCGCTTTCACAGGAAGCAATGGCGCC 402
 DB 1257 GYgGlyAlaArgPheGlyLeuSerThrGluAlaAsnGlyAla 1290

RESULT 9

P72197 PRELIMINARY; PRT; 1723 AA.
 AC P72197:
 DT 01-FEB-1997 (TREMBLrel. 02, Created)
 DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Lys-gingipain.
 GN KGP.
 OS Porphyromonas gingivalis (Bacteroides gingivalis).
 OC Bacteria; CFb group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
 OC Porphyromonas.
 OX NCBI_Taxid=837;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA Pavloff N., Pemberton P.A., Potempa J., Chen W.-C.A., Pike R.N.,
 RA Prochazka V., Kiefer M.C., Travis J., Barr P.J.;
 RT "Molecular cloning and characterization of Porphyromonas gingivalis
 RT Lys-gingipain.";
 RL Submitted (Apr-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U54691; AAA99810.1; -.
 DR MEROPS: C25.002; -.
 DR InterPro: IPR000977; DNA_Ligase.
 DR InterPro: IPR001769; Peptidase_C25.
 DR Pfam: PF01364; Peptidase_C25.3.
 DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
 SO SEQUENCE 1723 AA; 186831 MW; 4508A7E50197CEBD CRC64;

Alignment Scores:

Pred. No.: 8.19e-62 Length: 1723
 Score: 711.00 Matches: 130
 Percent Similarity: 99.25% Conservative: 3
 Best Local Similarity: 97.01% Mismatches: 1
 Query Match: 97.26% Indels: 0
 DB: 2 Gaps: 0

US-09-980-370-5 (1-402) x P72197 (1-1723)

OY 1 GCAGACTTCACGGAACGCTTCGAGTCTTCTACTATGAGAGGACCGCAATGGACT 60
 DB 1139 AlaAspPheThrGluThrPheGluSerSerThrHisGlyGluAlaProAlaGluThrPthr 1158
 OY 61 ACTATGATGCGGATGGGATGGAGGTTGGCTCTCTCTCTCTCGCAATGGAC 120
 DB 1159 ThrIleAspAlaAspGlyAspGlyGlnGlyIlePheLysLeuSerSerGlyGlnLeuAsp 1178

```

Db      1176 ThrIleaspAlaaspGlyaspGly6InglyTrpleucysLeuSerSerGly6Inleuasp 1195
QY      121  TGGCTCAGCTCATGCGCGGACCAACGAGTACGCTCTTCTCATGCAATGAATGGCT 180
Db      1196 TrpleuthrAlaHisGlyGlyTrhrasnValValAlaSerPheSerTrprasnGlyMetAla 1215
QY      181  TTGAATCCGTAACTATCTCATCTCAAAAGATGTTTACAGCGGCAACGAGTAAGTAC 240
Db      1216 LeuasnProaspasnTyrlleuIleSerLysaspValThrGlyAlaThrLysValLysTyr 1235
QY      241  TACTATCCAGTCACGACGAGCTTTCCCGGAGTACTACTAGCGGTATGATCTCCAGAGAG 300
Db      1236 TyrlTyrlAlaValAlaasnAspGlyPheProGlyaspHisTyrlAlaValMetIleSerLysThr 1255
QY      301  GGCACGACGACGCGGAGCTTACGCTTGTTCGAAAGAACCCCTAACGGAATAAATAAG 360
Db      1256 GlyThrasnAlaGlyaspPheThrValValPheGluGluThrProasnGlyIleasnLys 1275
QY      361  GCGCGAGACGAGATTCGGTCTTTCACAGCAAGCAATGCGGCC 402
Db      1276 GlyGlyAlaArgPheGlyLeuSerThrGluAlaaspGlyAla 1289

```

RESULT 10

```

ID      P96967      PRELIMINARY:      PRT: 1358 AA.
AC      P96967:
DT      01-MAY-1997 (TREMBLrel. 03, Created)
DT      01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DT      01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE      Hemagglutinin.
OS      Porphyromonas gingivalis (Bacteroides gingivalis).
OC      Bacteria; CFb group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
OC      Porphyromonas.
OX      NCBI_TaxID=837;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=381;
RA      Han N., Lepine G., Whitlock J., Wojciechowski L., Progulskie-Fox A.;
RT      "Cloning, sequencing and characterization of hagd, a member of the
RT      Harep multigene family in Porphyromonas gingivalis.";
RL      Submitted (Aug-1996) to the EMBL/GenBank/DBJ databases.
DR      EMBL: U68468; AAB49691.1;
DR      InterPro: IPR000977; DNA_Ligase.
DR      InterPro: IPR001769; Peptidase_C25.
DR      Pfam: PF01364; Peptidase_C25; 3.
DR      PROSITE: PS00697; DNA_LIGASE_A1;
SQ      SEQUENCE 1358 AA; 147102 MW; 47FCA0B25B06DE88 CRC64;

```

```

Alignment Scores:
Pred. No.:      1,24e-61      Length:      1358
Score:          709.00      Matches:      130
Percent Similarity: 98.51%      Conservative: 2
Best Local Similarity: 97.01%      Mismatches: 2
Query Match:    96.99%      Indels:      0
DB:             2          Gaps:      0

```

US-09-980-370-5 (1-402) x P96967 (1-1358)

```

QY      1  GCAGACTTCACGAAACGTTGAGTCTTCTACTCATGAGAGGACACACCGGAAATGCACT 60
Db      791  AlaaspPheThrGluThrPheGluSerSerThrHisGlyAlaProAlaGluTrpThr 810
QY      61  ACTATGATGCCGATGGCGATGGTGGAGGGTGGCTGTGCTGCTCCGCAATGGAGC 120
Db      811  ThrIleaspAlaaspGlyaspGlyInaspTrpleucysLeuSerSerGlyGlnleuasp 830
QY      121  TGGCTCAGCTCATGCGCGGACCAACGAGTACGCTCTTCTCATGGAATGAATGGCT 180
Db      831  TrpleuthrAlaHisGlyGlyTrhrasnValValAlaSerPheSerTrprasnGlyMetAla 850
QY      181  TTGAATCCGTAACTATCTCATCTCAAAAGATGTTTACAGCGGCAACGAGTAAGTAC 240

```

```

Db      851  LeuasnProaspasnTyrlleuIleSerLysaspValThrGlyAlaThrLysValLysTyr 870
QY      241  TACTATCCAGTCACGACGAGCTTTCCCGGAGTACTACTATGCGGTGATATCTCCAGACG 300
Db      871  TyrlTyrlAlaValAlaasnAspGlyPheProGlyaspHisTyrlAlaValMetIleSerLysThr 890
QY      301  GGCACGACGACCGGAGCTTACGCTTGTTCGAAAGAACCCCTAACGGAATAAATAAG 360
Db      891  GlyThrasnAlaGlyaspPheThrValValPheGluGluThrProasnGlyIleasnLys 910
QY      361  GCGCGACGACGAGATTCGGTCTTTCACAGCAAGCAATGCGGCC 402
Db      911  GlyGlyAlaArgPheGlyLeuSerThrGluAlaasnGlyAla 924

```

RESULT 11

```

ID      P72194      PRELIMINARY:      PRT: 1723 AA.
AC      P72194:
DT      01-FEB-1997 (TREMBLrel. 02, Created)
DT      01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT      01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE      Lys-gingipain.
OS      Porphyromonas gingivalis (Bacteroides gingivalis).
OC      Bacteria; CFb group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
OC      Porphyromonas.
OX      NCBI_TaxID=837;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=381;
RA      MEDLINE=97044756; PubMed=8889827;
RX      Okamoto K., Kadowaki T., Nakayama K., Yamamoto K.;
RT      "Cloning and sequencing of the gene encoding a novel lysine-specific
RT      cysteine proteinase (Lys-gingipain) in Porphyromonas gingivalis:
RT      structural relationship with the arginine-specific cysteine proteinase
RT      (Arg-gingipain)";
RL      J. Biochem. 120:398-406(1996).
DR      EMBL: D83258; BAA11870.1;
DR      MEROPS: C25.002;
DR      InterPro: IPR000977; DNA_Ligase.
DR      InterPro: IPR001769; Peptidase_C25.
DR      Pfam: PF01364; Peptidase_C25; 3.
DR      PROSITE: PS00697; DNA_LIGASE_A1;
SQ      SEQUENCE 1723 AA; 187261 MW; 5628963D251493EB CRC64;

```

```

Alignment Scores:
Pred. No.:      1,29e-61      Length:      1723
Score:          709.00      Matches:      130
Percent Similarity: 98.51%      Conservative: 2
Best Local Similarity: 97.01%      Mismatches: 2
Query Match:    96.99%      Indels:      0
DB:             2          Gaps:      0

```

US-09-980-370-5 (1-402) x P72194 (1-1723)

```

QY      1  GCAGACTTCACGAAACGTTGAGTCTTCTACTCATGAGAGGACACACCGGAAATGCACT 60
Db      1156  AlaaspPheThrGluThrPheGluSerSerThrHisGlyAlaProAlaGluTrpThr 1175
QY      61  ACTATGATGCCGATGGCGATGGTGGAGGGTGGCTGTGCTGCTCCGCAATGGAGC 120
Db      1176  ThrIleaspAlaaspGlyaspGlyInaspTrpleucysLeuSerSerGlyGlnleuasp 1195
QY      121  TGGCTCAGCTCATGCGCGGACCAACGAGTACGCTCTTCTCATGGAATGAATGGCT 180
Db      1196  TrpleuthrAlaHisGlyGlyTrhrasnValValAlaSerPheSerTrprasnGlyMetAla 1215
QY      181  TTGAATCCGTAACTATCTCATCTCAAAAGATGTTTACAGCGGCAACGAGTAAGTAC 240
Db      1216  LeuasnProaspasnTyrlleuIleSerLysaspValThrGlyAlaThrLysValLysTyr 1235
QY      241  TACTATCCAGTCACGACGAGTTCGCGGAGTACTACTATGCGGTGATATCTCCAGACG 300

```

```

Db      1236 TyrTyrAlaValaValaAspGlyPheProGlyAspHisTyrAlaValaMetIleSerLysThr 1255
QY      301 GGCACGAGACCGCGAGACTTCACGCTGTTTCGAGAAACGCGCTTAACGGAATAAATAG 360
Db      1256 GlyThrAsnAlaGlyAspPheThrValaIlePheGluGluThrProAsnGlyIleAsnLys 1275
QY      361 GCGGAGCAAGATTCGCTTCACAGCAAGCAATGGCGCC 402
Db      1276 GlyGlyAlaIaArgPheGlyLeuSerThrGluAlaAsnGlyAla 1289

RESULT 12
051838      PRELIMINARY:      PRT: 1706 AA.
ID 051838:
AC 051838:
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Protease precursor.
GN PRP1.
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria: CFb group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
OC Porphyromonas.
OX NCBI_TaxID=837;
  (1)
RP SEQUENCE FROM N.A.
RC STRAIN=W50.
RX MEDLINE=96071894; PubMed=7591131;
RA Aduse-Opoku J., Muir J., Slaney J.M., Rangarajan M., Curtis M.A.;
RT "Characterization, genetic analysis, and expression of a protease
  antigen (PzPRI) of Porphyromonas gingivalis W50."
RT Infect. Immun. 63:4744-4754(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=W50.
RA Rangarajan M., Aduse-Opoku J., Slaney J.M., Young K.A., Curtis M.A.;
RT "The PRP1 and the pR2 arginine-specific protease genes of
  Porphyromonas gingivalis W50 produce five biochemically distinct
  enzymes."
RT Mol. Microbiol. 23:0-0(1997).
DR EMBL: X82680; CA57997.1; -.
DR HSSP: P95493; ICVR.
DR MEROPS: C25.001; -.
DR InterPro: IPR000977; DNA_Ligase.
DR InterPro: IPR001769; Peptidase.C25.
DR Pfam: PF01364; Peptidase.C25.3.
DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
KW SIGNAL.
FT SIGNAL. 1 23 POTENTIAL.
FT CHAIN 228 719 ALPHA-PROTEASE.
FT CHAIN 720 1262 BETA-ADHESIN.
SQ SEQUENCE 1706 AA; 185705 MW; 0E56DCD87FDA8CDD CRC64;

Alignment Scores:
Pred. No.: 2,55e-61 Length: 1706
Score: 706.00 Matches: 130
Percent Similarity: 98.51% Conservative: 2
Best Local Similarity: 97.01% Mismatches: 2
Query Match: 96.58% Indels: 0
Db: 2 Gaps: 0

US-09-980-370-5 (1-402) x 051838 (1-1706)
QY 1 GCAGACTTCACGGAACGTTGCGAGTCTTCACTCATGAGAGCAGCAGCGGAATGACT 60
Db 1139 AlaAspPheThrGluThrPheGluSerSerThrHisGlyIuAlaThrAlaGluTTPThr 1158
QY 61 ACTATGATGCCGATGCCGATGAGAGGTTGGCTGTCTGTTCCGACAATTGGAC 120
Db 1159 ThrIleAspAlaAspGlyAspGlyGlnGlyTyrPleucysLeuSerSerGlyGlnLeuAsp 1178
QY 121 TGGCTCACACCTCATGGCGGACCAACGATAGACTCTTTCATGGAATGGAATGGCT 180

```

```

Db      1179 TrpLeuThrAlaHisGlyGlyThrAsnValaValaSerSerPheSerTrpAsnGlyMetAla 1198
QY      181 TTGAATCCGATACATATCTCATTCATAAGATGTTACAGCCCAACGAAGTAAGTAC 240
Db      1199 LeuAsnProAspAsnTyrLeuIleSerLysAspAlaThrGlyAlaThrLysValLysTyr 1218
QY      241 TACTATCCATCAACGACGCGTTTCCCGGATCCACTATGCGTGATGATCTCCAGACG 300
Db      1219 TyrTyrAlaValaValaAspGlyPheProGlyAspHisTyrAlaValaMetIleSerLysThr 1238
QY      301 GGCACGAGACCGCGAGACTTCACGCTGTTTCGAGAAACGCGCTTAACGGAATAAATAG 360
Db      1239 GlyThrAsnAlaGlyAspPheThrValaIlePheGluGluThrProAsnGlyIleAsnLys 1258
QY      361 GCGGAGCAAGATTCGCTTCACAGCAAGCAATGGCGCC 402
Db      1259 GlyGlyAlaIaArgPheGlyLeuSerThrGluAlaAspGlyAla 1272

RESULT 13
09F4J0      PRELIMINARY:      PRT: 925 AA.
ID 09F4J0:
AC 09F4J0:
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Putative outer membrane protein PG57.
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria: CFb group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
OC Porphyromonas.
OX NCBI_TaxID=837;
  (1)
RP SEQUENCE FROM N.A.
RC STRAIN=W50.
RA Ross B.C., Barr I., Patterson M., Agius C., Rothel L., Margetts M.,
  Hocking D., Webb E.;
RT "P. gingivalis polypeptides and nucleic acids."
RT Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=W50.
RA Ross B.C., Czajkowski L., Hocking D., Margetts M., Webb E., Rothel L.,
  Patterson M., Agius C., Camuglia S., Reynolds E., Barr I.G.;
RT "Identification of vaccine candidates from genomic analysis of
  Porphyromonas gingivalis."
RT Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY007522; AAC24228.1; -.
DR InterPro: IPR003961; FN_III.
DR SMART: SM00060; FN3; 3.
SQ SEQUENCE 925 AA; 103632 MW; 5F2198D6914DAE8 CRC64;

Alignment Scores:
Pred. No.: 8,25e-12 Length: 925
Score: 205.00 Matches: 55
Percent Similarity: 47.26% Conservative: 14
Best Local Similarity: 37.67% Mismatches: 43
Query Match: 28.04% Indels: 34
Db: 2 Gaps: 7

US-09-980-370-5 (1-402) x 09F4J0 (1-925)
QY 1 GCAGACTTCACGGAACGTTGCGAGTCTTCACT----- 33
Db 288 AlaValTyrAspGlySerIleGluSerSerThrValaLysGlyThrIleHisTyrAlaThr 307
QY 34 -----CATGAGAGCGACCGACCGGAATGAGACTACTATC 66
Db 308 AspAlaIleLeuTyrGluAsnPheGluAsnGlyProValProAsnGlyTyrPleuValIle 327
QY 67 GATCCGCGATGGCGATGGTGAGCGTTGGCTGTCTGTTCCGAGACAA----- 114
Db 328 AspAlaAspGlyAspGlyPheSerTrp-----GlyHisTyrLeuAsnAla 342
QY 115 TTGGACTGGCTCACAGCTCATGGCGGACCAACGATAGACTCTTTCATGCG---AAT 171

```

```

Db      343 TYRAspAlaPheProGlyHisAsnGlyGlyHisCysSerLeuSerAlaSerTyrValPro 362
QY      172 GGAATGGCT---TTGANTCCTGATTAATCTATCTATCTCAAGATGTTTACGGCGCAACG 228
Db      363 GylIleGlyProValThrProAspAsnTyrLeuIleThrProLysValGluGlyAlaLys 382
QY      229 AAGGTAAATGACTACTATCCAGTCAACGAC---GGTTTCCGGGATGACTATGCGGTG 285
Db      383 ArgValLysTyrTrpValSerThrGlnAspAlaAsnTrpAlaLagIuHisTyrAlaVal 402
QY      286 ATGATCTCCAAGACGGCGACGAGACTTCAGGCTTTTTCGAAGAACG--- 342
Db      403 MetAlaSerThrThrGlyThrAlaValGlyAspPheValIleLeuPheGluThrMet 422
QY      343 -----CCTAAGCA 351
Db      423 ThrAlaLysProThrGly 428

```

RESULT 14

```

O9KIB3 PRELIMINARY: PRT: 312 AA.
AC 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE Hypothetical outer membrane protein PG27.
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; CF8 group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
NCBI_TaxID=837;

```

```

RN 11)
RP SEQUENCE FROM N.A.
RC STRAIN=50;
RA Rose B., Barr I., Patterson M., Agius C., Rothel L., Margetts M.,
RT Hocking D., Webb E.,
RL "P. gingivalis polypeptides and nucleic acids.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF237555; AAF81413.1; -
DR InterPro: IPR001769; Peptidase_C25.
DR Pfam: PF01364; Peptidase_C25: 2.
SQ SEQUENCE 312 AA; 34592 MW; 0D5792C9643A25F5 CRC64;

```

Alignment Scores:

Pred. No.:	Length:	Score:	Matches:	Conservative:	Best Local Similarity:	Mismatches:	Indels:	Gaps:
2	2.07e-09	180.00	50	55.46%	42.02%	35	18	7
2	2.07e-09	180.00	50	55.46%	42.02%	35	18	7

US-09-980-370-5 (1-402) x O9KIB3 (1-312)

```

QY      10 ACGAAGAGTTCGAGTCTTCTACTGATGAGAGACGACGCAATGAGTACTATGAT 69
Db      5 SerGluSerPheGluSerGly-----IleProAlaIleTrpLysThrIleAsp 20
QY      70 GCCATGGCGATGATGAGGCTTGGCTTGTCTGTCTCCGACAATTGAGACTGGTCACA 129
Db      21 AlaAspGlyAspGlyTyrLysAsnTrpMetHisLeuThrAsn-----PheThr 35
QY      130 GCTATGCGCGACCAACAGTAAAGTCTTTCATGGAATGA-----ATGCGTTTG 183
Db      36 GlyIleSerGlyLeu---CysValIleSerAlaSerTyrIleGlyValGlyAlaLeu 54
QY      184 AATCGTGAATCACTATCTCAAGAGATGTTACAGCGGCAACG-----AAG 231

```

```

Db      55 ThrProAspAsnTyrIleuIleThrProGluLeuLysLeuProThrAspAlaLeuValGlu 74
QY      232 GTAAGACTACTATTCACAGTCAACGAC---GGTTTCCGGGATCACTATGCGGTATG 288
Db      75 IleIleTyrTrpValCysThrGlnAspLeuThrAlaProSerGluHisTyrAlaValTyr 94
QY      289 ATCTCCAGACGGCGACGACGACCGGAGACTTC---ACGTTTTCGAAGAACG 342
Db      95 SerSerSerThrGlyAsnAsnAlaAlaAspPheValAsnLeuLeuTyrGluGluThr 113

```

RESULT 15

```

O9JZ15 PRELIMINARY: PRT: 218 AA.
AC 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE Isomerase, putative.
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; Beta subdivision; Neisseriaceae; Neisseria.
NCBI_TaxID=491;

```

```

RN 11)
RP SEQUENCE FROM N.A.
RC STRAIN=MC58 / SEROGROUP B;
RX MEDLINE=2015755; PubMed=10710307;
RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA Nelson W.C., Gwinn M.L., Deboy R., Peterson J.D., Hickey E.K.,
RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA Mason T., Ciecko A., Parksey D.S., Blair E., Clifton H., Clark E.B.,
RA Cotton M.D., Ulterback T.R., Khouri H., Qin H., Vamathevan J.,
RA Gill J., Scarlato V., Maignani V., Pizza M., Grand G., Sun L.,
RA Smith H.O., Fraser C.W., Moxon E.R., Kapunoli R., Venter J.C.,
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
MC58.";
RL Science 287:1809-1815(2000).
DR EMBL: AE002482; AAF41713.1; -
DR TIGR: NMB1338; -
DR InterPro: IPR002529; FAA_hydrolase.
DR Pfam: PF01557; FAA_hydrolase; 1.
KW Complete proteome.
SQ SEQUENCE 218 AA; 23058 MW; BA531FCD9C010BD9 CRC64;

```

Alignment Scores:

Pred. No.:	Length:	Score:	Matches:	Conservative:	Best Local Similarity:	Mismatches:	Indels:	Gaps:
1	1.39	90.50	38	36.36%	26.57%	48	43	6
1	1.39	90.50	38	36.36%	26.57%	48	43	6

US-09-980-370-5 (1-402) x O9JZ15 (1-218)

```

QY      67 GATCCCGATGCGATGAGGAT----- 90
Db      77 AspSerAspGlyThrGluGluGlyAspIleLeuGlyCysValAlaGlyTyrGlyVal 96
QY      91 -----TGCGTCTGTCTGTCTTCCGACAATTGAGACTGG 123
Db      97 GlyLeuAspLeuThrAlaArgAspIleGlnCysArgLeuLysGlyLeuGlyLeuProThr 116
QY      124 CTCACAGCTCATGGC-----GGACCAACAGTAAAGTCTTTCATGGAATGGAATG 177
Db      117 LeuLysAlaLysGlyPheArgHisSerAlaCysValSerAspPheAlaAlaIleGlyArg 136
QY      178 GCTTGAATCCTGATTAATATATCTCATGCAAGATGTTACAGCGCAACGAGTAAG 237
Db      137 IleGlyAsnProGluLysValLeuPheSerLeuLysGlnAsnGlyValLeuLysGlnArg 156
QY      238 -----TACTATATTCACAGTCAACGAGGTTTCCGGGAGTCACTATGCG 282
Db      157 GlyAspThrGlyLeuMetIleTyrProIleArgGlu----- 168

```

```

OY      283 GTGATGATCTCCAGACG-----GCCACGAACCGCGAGACTTCACGGTTGTT 330
          ::::: |||
Db      169 ---IleLeuH1sLysLeuAla1aAspTyrGlyLeuGlyLysGlyasp-----LeuVal 185
          ::::: |||
OY      331 TTGGAAGAAACGCCCTAACGGAATAAATAAAGGCGGACCAAGATTCCGTCCTTTCCACGGAA 390
          ||| ||| ::::: |||
Db      186 PheThrGlyThrProSerGlyValGlyAla1IleGlyAlaGlyAspAsnLeuAlaLeuGlu 205
          ::::: |||
OY      391 GCCAATGGC 399
          ::::: |||
Db      206 LeuAspGly 208

```

Search completed: June 26, 2003, 00:44:14
 Job time : 44.5 secs

!

Thu Jun 26 11:58:13 2003

us-09-980-370-6.frag

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 26, 2003, 00:44:59 ; Search time 47.2941 Seconds
(without alignments)
377.543 Million cell updates/sec

Title: US-09-980-370-6
Perfect score: 730
Sequence: 1 ADPTEPESSTHGEAPAEWT.....PNGINKGARGLESTENGCA 134

Scoring table:
BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database:

1: A_Geneseq_101002.*
2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
20: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
21: /SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
24: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	730	100.0	134	AA19217	Peptide used in th
2	716	98.1	1687	AA19603	P. gingivalis haem
3	716	98.1	1687	AA19603	Haemagglutinin pro
4	716	98.1	1704	AA19603	Arg-gingipain 2 pt
5	716	98.1	1704	AA19603	Arg-gingipain 2 pt
6	716	98.1	1704	AA19603	Arg-gingipain 2 pt
7	716	98.1	1704	AA19603	Arg-gingipain 2 pt
8	716	98.1	1704	AA19603	Arg-gingipain 2 pt
9	716	98.1	1704	AA19603	Arg-gingipain 2 pt
10	716	98.1	1704	AA19603	Arg-gingipain 2 pt

11	714	97.8	1706	AA19603	P. gingivalis haem
12	711	97.4	439	AA19603	Haemagglutinin pro
13	711	97.4	439	AA19603	Arg-gingipain 2 pt
14	711	97.4	450	AA19603	Arg-gingipain 2 pt
15	711	97.4	450	AA19603	Arg-gingipain 2 pt
16	711	97.4	450	AA19603	Arg-gingipain 2 pt
17	711	97.4	450	AA19603	Arg-gingipain 2 pt
18	709	97.1	1087	AA19603	Arg-gingipain 2 pt
19	709	97.1	1087	AA19603	Arg-gingipain 2 pt
20	709	97.1	1087	AA19603	Arg-gingipain 2 pt
21	709	97.1	1087	AA19603	Arg-gingipain 2 pt
22	709	97.1	1087	AA19603	Arg-gingipain 2 pt
23	709	97.1	1087	AA19603	Arg-gingipain 2 pt
24	709	97.1	1087	AA19603	Arg-gingipain 2 pt
25	709	97.1	1087	AA19603	Arg-gingipain 2 pt
26	709	97.1	1087	AA19603	Arg-gingipain 2 pt
27	709	97.1	1087	AA19603	Arg-gingipain 2 pt
28	709	97.1	1087	AA19603	Arg-gingipain 2 pt
29	709	97.1	1087	AA19603	Arg-gingipain 2 pt
30	709	97.1	1087	AA19603	Arg-gingipain 2 pt
31	709	97.1	1087	AA19603	Arg-gingipain 2 pt
32	709	97.1	1087	AA19603	Arg-gingipain 2 pt
33	709	97.1	1087	AA19603	Arg-gingipain 2 pt
34	709	97.1	1087	AA19603	Arg-gingipain 2 pt
35	709	97.1	1087	AA19603	Arg-gingipain 2 pt
36	709	97.1	1087	AA19603	Arg-gingipain 2 pt
37	709	97.1	1087	AA19603	Arg-gingipain 2 pt
38	709	97.1	1087	AA19603	Arg-gingipain 2 pt
39	709	97.1	1087	AA19603	Arg-gingipain 2 pt
40	709	97.1	1087	AA19603	Arg-gingipain 2 pt
41	709	97.1	1087	AA19603	Arg-gingipain 2 pt
42	709	97.1	1087	AA19603	Arg-gingipain 2 pt
43	709	97.1	1087	AA19603	Arg-gingipain 2 pt
44	709	97.1	1087	AA19603	Arg-gingipain 2 pt
45	709	97.1	1087	AA19603	Arg-gingipain 2 pt

ALIGNMENTS

RESULT 1	AA19217	standard; protein: 134 AA.
ID	AA19217	
AC	AA19217	
DT	13-MAR-2001	(first entry)
DE	Peptide used in the invention.	
XX	HA2 domain; porphyrin; periodontal; pulmonary; vaginal; urethral;	
KW	Infection.	
OS	Unidentified.	
PN	WO200072875-A1.	
XX	07-DEC-2000.	
XX	26-MAY-2000: 2000MO-AU00599.	
XX	28-MAY-1999: 99AU-0000652.	
XX	(UNSY) UNIV SYDNEY.	
XX	Colliger CA, Hunter N, De Carlo AA:	
XX	WPI: 2001-080424/09.	
PT	Treating microbial infection in environment containing porphyrin, by	
PT	administering a HA-2 antagonist	
PS	Claim 9; Page 98-99; 102pp; English.	

xx The present invention relates to prophylaxis or treatment of infection
 CC by a microorganism in a biological environment comprising iron, heme
 CC or porphyrin. The treatment involves administering to the environment,
 CC an agent which antagonizes the interaction between a molecule derived
 CC from the microorganism and having an h4z domain, and an h4z-binding
 CC molecule on a porphyrin containing molecule present in the environment.
 CC The invention also relates to a method of a medicine for the prophylaxis and
 CC treatment of the heme deficiency anemia, peridontal, periodontal or foot disease
 CC resulting from porphyromonas gingivalis infection or infection by a
 CC related microorganism.

SO Sequence 134 AA:

Query Match 100.0%; Score 730; DB 22; Length 134;
 Best Local Similarity 100.0%; Pred. No. 7.9e-68;
 Matches 134; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ADFTTFESSSTHGEAPAEWTTIDADGDEGMLCLSSGOLDMLTAHGTNNVVSFSNMGNA 60
 DB 1 ADFTTFESSSTHGEAPAEWTTIDADGDEGMLCLSSGOLDMLTAHGTNNVVSFSNMGNA 60
 OY 61 LNPONTLSKDVGTATKVKYYPVNDGFPDGHVAVMISKTGTNAGDFTVVEETPNCINK 120
 DB 61 LNPONTLSKDVGTATKVKYYPVNDGFPDGHVAVMISKTGTNAGDFTVVEETPNCINK 120
 OY 121 GCARFGLSTEANGA 134
 DB 121 GCARFGLSTEANGA 134

RESULT 2
 AAR96033 standard: Protein: 1687 AA.
 ID AAR96033:
 AC AAR96033:
 DT 04-SEP-1996 (first entry)
 DE P. gingivalis haemagglutinin hage.
 KW Haemagglutinin; hage; periodontal disease; vaccine; antibody.
 OS Porphyromonas gingivalis strain FDC381.
 PN W09617936-A2.
 PD 13-JUN-1996.
 PR 11-DEC-1995; 95MO-US16108.
 PR 09-DEC-1994; 94US-0353485.
 PA (UABR-) UAB RES FOUND.
 PA (UYFL) UNIV FLORIDA.
 PI Han N, Lantz M, Lepline G, Patell JM, Proguliske-Fox A;
 PI Tumaseorn S;
 PI WPI: 1996-287181/29.
 DR N-PSDB: AAT30656.
 PT Porphyromonas gingivalis genes and proteins - used in the detection
 PT and vaccination against periodontal disease
 PS Claim 5; Page 138-143; 153pp; English.
 CC P. gingivalis 381 haemagglutinin hage (AAR96033) was identified as
 CC the product of a gene (AAT30656) identified in P. gingivalis 318
 CC genomic DNA. The haemagglutinin can be used from transformed
 CC host cells and used as a vaccine to protect humans or animals against
 CC periodontal disease. Expression in *Salmonella* cells allows prodn. of
 CC a live vaccine. The haemagglutinin can also be used to detect the

CC presence of anti-P. gingivalis antibodies and to raise monoclonal
 CC antibodies for diagnostic appln.
 CC Sequence 1687 AA:
 SO Sequence 1687 AA:

Query Match 98.1%; Score 716; DB 17; Length 1687;
 Best Local Similarity 97.8%; Pred. No. 7.1e-65;
 Matches 131; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 ADFTTFESSSTHGEAPAEWTTIDADGDEGMLCLSSGOLDMLTAHGTNNVVSFSNMGNA 60
 DB 1120 ADFTTFESSSTHGEAPAEWTTIDADGDEGMLCLSSGOLDMLTAHGTNNVVSFSNMGNA 1179
 OY 61 LNPONTLSKDVGTATKVKYYPVNDGFPDGHVAVMISKTGTNAGDFTVVEETPNCINK 120
 DB 1180 LNPONTLSKDVGTATKVKYYPVNDGFPDGHVAVMISKTGTNAGDFTVVEETPNCINK 1239
 OY 121 GCARFGLSTEANGA 134
 DB 1240 GCARFGLSTEANGA 1253

RESULT 3
 AAM69495 standard: Protein: 1687 AA.
 ID AAM69495:
 AC AAM69495:
 DT 22-DEC-1998 (first entry)
 DE Haemagglutinin protein hage.
 KW Haemagglutinin protein; periodontal disease; vaccine; hage.
 OS Porphyromonas gingivalis.
 PN US5824791-A.
 PD 20-OCT-1998.
 PR 11-DEC-1995; 95US-0570311.
 PR 11-DEC-1995; 95US-0570311.
 PR 08-SEP-1988; 88US-0241640.
 PR 25-JAN-1991; 91US-0647119.
 PR 09-DEC-1994; 94US-0353485.
 PA (UABR-) UAB RES FOUND.
 PA (UYFL) UNIV FLORIDA.
 PI Han N, Lantz M, Lepline G, Patell JM, Proguliske-Fox A;
 PI Tumaseorn S;
 PI WPI: 1998-582627/49.
 DR N-PSDB: AAV58881.
 PT Isolated Porphyromonas gingivalis genes - encoding haemagglutinin
 PT and/or protease poly:peptide(s)
 PS Claim 1; Column 167-182; 101pp; English.
 CC This sequence is encoded by a Porphyromonas gingivalis gene of the
 CC invention. This sequence represents the hage haemagglutinin protein. The
 CC polypeptides are used to produce antibodies to organisms associated with
 CC periodontal disease. The antibodies are also used in purification and
 CC identification procedures. The genes and polypeptides are used as
 CC vaccines against periodontal disease.

SO Sequence 1687 AA:
 Query Match 98.1%; Score 716; DB 19; Length 1687;
 Best Local Similarity 97.8%; Pred. No. 7.1e-65;
 Matches 131; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ADFTETFEESSTHGCAPEAWMTTIDADGEGWCLSSGOLDMLTAHGCTNVVSSFSWNGMA 60
DB 1120 ADFTETFEESSTHGCAPEAWMTTIDADGEGWCLSSGOLDMLTAHGCTNVVSSFSWNGMA 1179
QY 61 LNPDNLTISKDYTGATKVKYKYYAVNDGFPDGHYAVMISKTGTNAGDFTVVEETPNCINK 120
DB 1180 LNPDNLTISKDYTGATKVKYKYYAVNDGFPDGHYAVMISKTGTNAGDFTVVEETPNCINK 1239
QY 121 GGARFGLSTEANGA 134
DB 1240 GGARFGLSTEANGA 1253

RESULT 4
AAR70188

ID AAR70188 standard; protein: 1704 AA.

AC AAR70188;

DT 21-SEP-1995 (first entry)

DE Arg-ginglipain-2 prepolyprotein.

KW Arg-ginglipain-2; gingivitis; periodontal disease; vaccine;
arginine-specific protease.

XX Porphyromonas gingivalis.

FH Key Location/Qualifiers

FT 228..719

FT /label= Protease

FT /note= "corresponds to Arg-ginglipain-1"

FT 720..1091

FT /label= Hemagglutinin

FT 1092..1429

FT /label= Hemagglutinin

FT 1430..1704

FT /label= Hemagglutinin

XX MO9507286-A.

PD 16-MAR-1995.

XX 09-SEP-1994; 94MO-US10283.

XX 24-JUN-1994; 94US-0265441.

XX 10-SEP-1993; 93US-0119361.

XX 21-OCT-1993; 93US-0141324.

XX (UYGE-) UNIV GEORGIA RES FOUND INC.

XX Barr PJ, Pavloff N, Potempa J, Travis J;

XX MPI; 1995-12373/16.

XX N-PSDB; AAO83489.

XX DNA encoding Arg-ginglipain proteins - used to develop prods. for

XX detection, treatment and prevention of periodontal disease

XX disclosure; Page 70-77; 89pp; English.

XX A low mol. wt. arginine-specific gingipain (AG-1) and high mol. wt.
XX AG (AG-2) were isolated from P. gingivalis strains H66 (ATCC
XX 33277) and M50 (ATCC 53973). The sequences of the proteins were
XX used to design PCR primers and probes to isolate AG DNA. Lambda
XX DASH and lambda ZAP libraries were screened with a probe based on
XX amino acids 11-22 of the AG protein to obtain DNA encoding AG-1
XX (AAO83484) and AG-2 (AAO83489). AG-2 is a prepolyprotein
XX incorporating AG-1.

SQ Sequence 1704 AA;

Query Match 98.1%; Score 716; DB 16; Length 1704;
Best Local Similarity 97.8%; Pred. No. 7,2e-65;
Matches 131; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ADFTETFEESSTHGCAPEAWMTTIDADGEGWCLSSGOLDMLTAHGCTNVVSSFSWNGMA 60
DB 1137 ADFTETFEESSTHGCAPEAWMTTIDADGEGWCLSSGOLDMLTAHGCTNVVSSFSWNGMA 1196
QY 61 LNPDNLTISKDYTGATKVKYKYYAVNDGFPDGHYAVMISKTGTNAGDFTVVEETPNCINK 120
DB 1197 LNPDNLTISKDYTGATKVKYKYYAVNDGFPDGHYAVMISKTGTNAGDFTVVEETPNCINK 1256
QY 121 GGARFGLSTEANGA 134
DB 1257 GGARFGLSTEANGA 1270

RESULT 5

ID AAW34843 standard; protein: 1704 AA.

AC AAW34843;

DT 03-JUN-1998 (first entry)

DE Arg-ginglipain high molecular weight prepolyprotein sequence.

KW Arg-specific gingipain protease; gingivitis; periodontal disease;
vaccine; infection.

XX Porphyromonas gingivalis.

FH Key Location/Qualifiers

FT 1..227

FT /note= "precursor protein"

XX MO9734629-A1.

XX 25-SEP-1997.

XX 21-MAR-1997; 97WO-US04635.

XX 22-MAR-1996; 96US-0013945.

XX (MORE-) MOREHOUSE SCHOOL MEDICINE.

XX (UYGE-) UNIV GEORGIA RES FOUND INC.

XX Genco CA, Potempa J, Travis J, Genco C;

XX MPI; 1997-479993/44.

XX N-PSDB; NAT93872.

XX Porphyromonas gingivalis Arg-specific gingipain protease peptide(s)

XX - useful for protecting animals and humans from gingivitis and

XX periodontal diseases

XX disclosure; Pages 68-73; 95pp; English.

XX The present sequence represents an arginine-specific protease of
XX Porphyromonas gingivalis. The following peptides, derived from Arg-
XX and Lys-specific high molecular weight proteases, offer protection
XX against infection: YTVVYRDGK IKEGTLATTE DGVATGNEH YGVEKYTAGS VSPKYC
XX (I): YTPVEEKONG RMIVIVAKKY (II): QLPPIPVAC VNGDFEFSMP CFEAALMRNO
XX (III): GEPNRPQPS NITATTGOK VTLKMDAPSTR (IV): GNEHYCEYK YTAGVSPKYC
XX KDVTG (V): RKEWNEPER YTPVEEKONG (VI): TFAGFEDTYR RMPMNEPER (VII):
XX KIEGTLATTE EEDG (IX): RDTGKIEGL TATVEEDGV ATGN (X): KIEGTLATTE
XX FEEDGATGN HEY (XI): KMDAPNGPN PNP PNPNGN PGTTLSE (XII): and
XX YTPVEEKONG RMIVIVAKKY (XIII). They are used in vaccines to protect
XX animals, including humans, from gingivitis and/or periodontal
XX diseases.

SQ Sequence 1704 AA;

```

Query Match          98.1%; Score 716; DB 18; Length 1704;
Best Local Similarity 97.8%; Pred. No. 7.2e-65;
Matches 131; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ADFTETFEESTHGEAPAEWTTIDADGGGWLCLSSGQLDMLTAHGCTNVVASFSGMGA 60
   |||||
Db 1137 ADFTETFEESTHGEAPAEWTTIDADGGGWLCLSSGQLDMLTAHGCTNVVASFSGMGA 1196
   |||||

QY 61 LMPDNVLLSKDVTGATKVKYYPVNDGFPGDHYAVMISKTGTNAGDFTVFEETPNGINK 120
   |||||
Db 1197 LMPDNVLLSKDVTGATKVKYYPVNDGFPGDHYAVMISKTGTNAGDFTVFEETPNGINK 1256
   |||||

QY 121 GGARFGLSTEANGA 134
   |||||
Db 1257 GGARFGLSTEANGA 1270
   |||||

RESULT 6
AAV67396
ID AAV67396 standard; Protein; 1704 AA.
XX AC AAV67396;
XX DT 25-APR-2000 (first entry)
XX DE Arg-gingipain-2 amino acid sequence.
XX KW Arginine specific proteinase; Arg-gingipain; gingipain-2; haemagglutinin;
XX KM immunogenic component; vaccine; inflammatory response; tissue damage;
XX OS Porphyromonas gingivalis.
XX FH Key Location/Qualifiers
   FT Region 229..719
   FT /note= "Amino acids 229-719 are specifically claimed"
   FT Region 720..1185
   FT /note= "Amino acids 720-1185 are specifically claimed"
XX PN US6017532-A.
XX PD 25-JAN-2000.
XX PF 08-NOV-1994; 94US-0336308.
XX PR 10-SEP-1993; 93US-0119361.
XX PR 24-JUN-1994; 94US-0265441.
XX PA (UYGE-) UNIV GEORGIA RES FOUND INC.
XX PI Potempa JS, Travis J;
XX DR MPI: 2000-136659/12.
XX DR N-PSDB: AAZ60181.
XX PT New Porphyromonas gingivalis arginine-specific protease preparation
XX PT useful for preparing vaccines against periodontal disease and for
XX PT screening for Arg-gingipain inhibitors
XX PS Claim 1: Column 29-42; 55pp; English.
XX XX This sequence represents a Porphyromonas gingivalis arginine-specific
XX CC proteinase known as Arg-gingipain/gingipain-2 amino acid sequence.
XX CC Gingipain-2 consists of a 50kD protease component non-covalently
XX CC associated with a 44kD haemagglutinin component. The proteinase is
XX CC stimulated by glycine containing peptides and glycine analogues. It is
XX CC inhibited by cysteine protease group specific inhibitors. The protease
XX CC preparation can be used in immunogenic compositions and vaccines against
XX CC inflammatory response and tissue damage caused by P. gingivalis in
XX CC periodontal disease. It can also be used to screen for agents that
XX CC modulate Arg-gingipain proteinase activity inhibitors.

```

```

SQ Sequence 1704 AA;
Query Match          98.1%; Score 716; DB 21; Length 1704;
Best Local Similarity 97.8%; Pred. No. 7.2e-65;
Matches 131; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ADFTETFEESTHGEAPAEWTTIDADGGGWLCLSSGQLDMLTAHGCTNVVASFSGMGA 60
   |||||
Db 1137 ADFTETFEESTHGEAPAEWTTIDADGGGWLCLSSGQLDMLTAHGCTNVVASFSGMGA 1196
   |||||

QY 61 LMPDNVLLSKDVTGATKVKYYPVNDGFPGDHYAVMISKTGTNAGDFTVFEETPNGINK 120
   |||||
Db 1197 LMPDNVLLSKDVTGATKVKYYPVNDGFPGDHYAVMISKTGTNAGDFTVFEETPNGINK 1256
   |||||

QY 121 GGARFGLSTEANGA 134
   |||||
Db 1257 GGARFGLSTEANGA 1270
   |||||

RESULT 7
AAU08938
ID AAU08938 standard; Protein; 1704 AA.
XX AC AAU08938;
XX DT 18-DEC-2001 (first entry)
XX DE P. gingivalis high molecular weight Arg-gingipain-2.
XX KW Periodontitis; antiinflammatory; Arg-gingipain-2; AG-2; Immunogen.
XX OS Porphyromonas gingivalis.
XX FH Key Location/Qualifiers
   FT Peptide 1..227
   FT /label= Prepro-peptide
   FT Region 670..674
   FT /label= Proteolytic_component
   FT Cleavage-site 719
   FT Protein 720..1091
   FT /label= HGP-44kDa
   FT /note= "Haemagglutinin protein component"
   FT Region 599..619
   FT /note= "Region of homology with cysteine proteases"
   FT Cleavage-site 1091
   FT Protein 1092..1429
   FT /label= HGP-17kDa
   FT Cleavage-site 1429
   FT /note= "Haemagglutinin protein component"
   FT Protein 1430..1704
   FT /label= HGP-27kDa
   FT /note= "Haemagglutinin protein component"
XX PN US6274718-B1.
XX PN 14-AUG-2001.
XX PD 25-JAN-2000; 2000US-0490931.
XX PF 24-JUN-1994; 94US-0265441.
XX PR 08-NOV-1994; 94US-0336308.
XX PR 10-SEP-1993; 93US-0119361.
XX PR 09-SEP-1994; 94WO-US10283.
XX PA (UYGE-) UNIV GEORGIA RES FOUND INC.
XX PI Travis J, Potempa JS, Barr PJ, Pavloff N;
XX DR MPI: 2001-588904/66.
XX DR N-PSDB: AAS15242.
XX PT New recombinant DNA molecule which encodes high molecular weight
XX PT (mature) Arg-gingipain protein, useful for immunisation against

```

PT Inflammation and tissue damage, comprises enzymatically active protease
 PT component and haemagglutinin component
 PS Claim 1: Column 29-41: 56pp: English.

XX The invention relates to a recombinant DNA molecule encoding high
 CC molecular weight (mature) Arg-gingipain (AG) protein, which has an
 CC enzymatically active protease component (AG-2) and a haemagglutinin
 CC component, from P. gingivalis. The nucleic acid is useful for producing
 CC mature Arg-gingipain protein. Immunogenic compositions comprising
 CC Arg-gingipain are useful for immunising animals including humans against
 CC inflammatory response and tissue damage caused by an archaeobacterium
 CC Porphyromonas gingivalis, which causes progressive periodontitis.
 CC Arg-gingipain is also useful for identifying agents that modulate
 CC Arg-gingipain protease activity, whether by acting on the protease
 CC itself or preventing the interaction of the protease with the protein
 CC in the gingival area, such as complement factors C3 or C5. The
 CC present sequence is Arg-gingipain-2.

XX Sequence 1704 AA:

Query Match 98.1%; Score 716; DB 22; Length 1704:
 Best Local Similarity 97.8%; Pred. No. 7.2e-65;
 Matches 131; Conservative 2; Mismatches 1; Indels 0; Gaps 0:

OY 1 ADFTTFESTSTHGEAPAEWTTIDADGGGWLCLSSGQLDWLTAGGTNNVSSFSMNMA 60
 DB 1137 ADFTTFESTSTHGEAPAEWTTIDADGGGWLCLSSGQLDWLTAGGTNNVSSFSMNMA 1196
 OY 61 LNPDNLTISKDVYGATKRVKYYYPVNDGFGDHYAVAMISKTGTNAGDFVVEEPTNGINK 120
 DB 1197 LNPDNLTISKDVYGATKRVKYYYPVNDGFGDHYAVAMISKTGTNAGDFVVEEPTNGINK 1256
 OY 121 GGARFGLSTENGA 134
 DB 1257 GGARFGLSTENGA 1270

OY 121 GGARFGLSTENGA 134
 DB 1257 GGARFGLSTENGA 1270

RESULT 8
 AAR96029 standard; Protein: 1732 AA.

XX AAR96029;
 DT 04-SEP-1996 (first entry)

DE P. gingivalis porphyrein.

KW Porphyrein: haemagglutinin; periodontal disease; vaccine; antibody.

XX Porphyromonas gingivalis strain W12.

OS Location/Qualifiers

FT Region 688..708

FT Region 887..952

FT Region 946..967

FT Region 985..1006

FT Region 1041..1100

FT Region 1341..1405

FT Region 1430..1451

FT Region 1488..1547

FT Region 1607..1650

FT Region /note="Pro-Asn repeat region type 2"

FT Region /note="Pro-Asn repeat region type 3"

FT Region /note="Pro-Asn repeat region type 4"

FT Region /note="Pro-Asn repeat region type 4"

FT Region /note="Pro-Asn repeat region type 4"

FT Region /note="Pro-Asn repeat region type 4"

FT Region /note="Pro-Asn repeat region type 4"

FT Region /note="Pro-Asn repeat region type 4"

PN M09617936-A2.

XX 13-JUN-1996.

XX 11-DEC-1995; 95MO-US16108.

XX 09-DEC-1994; 94US-0353485.

XX (UABR-) UAB RES FOUND.

XX (UYFL) UNIV FLORIDA.

XX Han N, Lantz M, Lepine G, Patil JM, Progulake-Fox A;

XX Tumasorn S;

XX WPI: 1996-287181/29.

XX N-PSDB; AAT30653.

XX Claim 5: Page 76-81: 153pp: English.

XX P. gingivalis W12 cysteine protease, porphyrein (AAR96029), was
 CC identified as the product of the prt gene (AAT30653) isolated from
 CC P. gingivalis W12 genomic DNA. The porphyrein shows homology to
 CC the haemagglutinins (see also AAR96028-28 and AAR96030-33) of P.
 CC gingivalis 318. It can be obtained from transformed host cells and
 CC used as a vaccine to protect humans or animals against periodontal
 CC disease. Expression in *Salmonella* cells allows production of a live
 CC vaccine. The porphyrein and haemagglutinins can also be used to
 CC detect the presence of anti-P. gingivalis antibodies and to raise
 CC monoclonal antibodies for diagnostic applications.

XX Sequence 1732 AA:

Query Match 97.9%; Score 715; DB 17; Length 1732:
 Best Local Similarity 97.8%; Pred. No. 9.4e-65;
 Matches 131; Conservative 2; Mismatches 1; Indels 0; Gaps 0:

OY 1 ADFTTFESTSTHGEAPAEWTTIDADGGGWLCLSSGQLDWLTAGGTNNVSSFSMNMA 60
 DB 1157 ADFTTFESTSTHGEAPAEWTTIDADGGGWLCLSSGQLDWLTAGGTNNVSSFSMNMA 1216
 OY 61 LNPDNLTISKDVYGATKRVKYYYPVNDGFGDHYAVAMISKTGTNAGDFVVEEPTNGINK 120
 DB 1217 LNPDNLTISKDVYGATKRVKYYYPVNDGFGDHYAVAMISKTGTNAGDFVVEEPTNGINK 1276
 OY 121 GGARFGLSTENGA 134
 DB 1277 GGARFGLSTENGA 1290

OY 121 GGARFGLSTENGA 134
 DB 1277 GGARFGLSTENGA 1290

RESULT 9
 AAM24787 standard; Protein: 1732 AA.

XX AAM24787;

XX 25-NOV-1997 (first entry)

DE PTK antigenic protein complex.

KW Periodontal disease; cell surface protein; chitin protease;

KW haemagglutinin; adhesin; therapy; diagnosis; vaccine; antigen.

OS Porphyromonas gingivalis strain W50.

FT Key Location/Qualifiers

FT Peptide 1..228

FT Cleavage-site /label="Pro-pro-peptide

FT Protein 229..737

FT /label="Prtk48

DE Prtr antigenic protein complex.

XX Periodontal disease; cell surface protein; thiol protease;

KM endopeptidase; Prtr: Prtr45; Prtr45; Prtr17; Prtr27;

KM haemagglutinin; adhesin; therapy; diagnosis; vaccine; antigen.

XX Porphyromonas gingivalis strain W50.

XX Key Location/Qualifiers

FT Peptide 1..227

FT /label= Pro-pro-peptide

FT Cleavage-site 227..228

FT Protein 228..719

FT /label= Prtr45

FT /note= "45 kDa Arg-specific thiol protease"

FT Cleavage-site 719..720

FT Protein 720..1138

FT /label= Prtr44

FT /note= "44 kDa adhesin"

FT Cleavage-site 1138..1139

FT Protein 1139..1273

FT /label= Prtr15

FT /note= "15 kDa adhesin"

FT Cleavage-site 1273..1274

FT Protein 1274..1431

FT /label= Prtr17

FT /note= "17 kDa adhesin"

FT Cleavage-site 1431..1432

FT Protein 1432..1706

FT /label= Prtr27

FT /note= "27 kDa adhesin"

XX W09716542-A1.

XX 09-MAY-1997.

XX 30-OCT-1996: 96WO-AU00673.

XX 30-OCT-1995: 95AU-0006275.

XX (UYME-) UNIV MELBOURNE.

XX (VICT-) VICTORIAN DAIRY IND AUTHORITY.

XX Bhogal PS, Reynolds EC, Slakeski N;

XX MPI: 1997-272112/24.

XX N-PSDB: AAT78850.

XX New antigenic protein complex from Porphyromonas gingivalis -

XX comparing Arg- and Lys- specific thiol endopeptidase(s), used in

XX the detection, prevention and treatment of periodontal disease

XX Example 1: Fig 8b: 68pp: English.

XX A Prtr-Prtr cell surface protein of Porphyromonas gingivalis (PG)

XX comprises a 300 kDa complex composed of a 45 kDa arginine-specific

XX thiol protease and 44, 15, 17 and 27 kDa adhesins encoded by the

XX Prtr gene (AAT78850), and a 146 kDa lysine-specific thiol protease

XX and 39, 15 and 44 kDa adhesins (see AAW24787) encoded by the Prtr

XX gene (AAT78851). A claimed antigenic complex comprises at least one

XX multimeric protein complex of Prtr and Prtr each containing at

XX least one adhesion domain, the complex having a mol. wt. of over 200

XX kDa, and preferably comprises all 9 proteins of the Prtr-Prtr

XX complex (see also AAW24780-85). It can be used in a claimed

XX composition to elicit an immune response directed against PG, and

XX in a claimed method of reducing the prospect of PG infection and/or

XX severity of disease. Antibodies directed against the complex are

XX claimed for use in treating PG infection. Unlike whole PG cells or

XX other previously prepared antigens based on fimbriae or the

XX capsule, the Prtr-Prtr complex or component parts are safe and

XX effective antigens.

XX Sequence 1706 AA:

Query Match 97.8%: Score 714: DB 18: Length 1706:

Best Local Similarity 97.8%: Pred No. 1.2e-64:

Matches 131: Conservative 2: Mismatches 1: Indels 0: Gaps 0:

QY 1 ADFTFEESSTHGEAPAEWTTIDADGEGWCLSSGQDLMTLAHGTNVVSSFSWNGMA 60

DB 1139 ADFTFEESSTHGEAPAEWTTIDADGEGWCLSSGQDLMTLAHGTNVVSSFSWNGMA 1198

QY 61 LNPDNVILSKDVGATVKYYPVNDGFPDGHVAVMISKTGTNAGDPVWFEEETPGINK 120

DB 1199 LNPDNVILSKDVGATVKYYPVNDGFPDGHVAVMISKTGTNAGDPVWFEEETPGINK 1258

QY 121 GGARFGLSTEANGA 134

DB 1259 GGARFGLSTEANGA 1272

RESULT 12

AAR96024

ID AAR96024 standard; Protein: 439 AA.

XX AAR96024:

AC AAR96024:

DT 04-SEP-1996 (first entry)

DE P. gingivalis haemagglutinin haga Harep4 product.

XX Haemagglutinin: haga: periodontal disease; vaccine; antibody;

KW Harep4.

XX Porphyromonas gingivalis strain 381.

OS W09617936-A2.

PN 13-JUN-1996.

PD 11-DEC-1995: 95WO-US16108.

PF 09-DEC-1994: 94US-0353485.

PR (UABR-) UAB RES FOUND.

XX (UFL) UNIV FLORIDA.

XX Han N, Lantz M, Leptine G, Patli JM, Proguliske-Fox A;

XX Tumwasorn S;

XX MPI: 1996-287181/29.

XX N-PSDB: AAT30648.

XX Porphyromonas gingivalis genes and proteins - used in the detection

XX and vaccination against periodontal disease

XX Claim 4: Page 114-115; 153pp: English.

XX Harep4 (AAR96024) is the product of the Harep4 repeat unit (AAT30648)

XX of the haga gene (AAT30654) of P. gingivalis 318. It forms part

XX of haemagglutinin haga (see also AAR96030). Harep4 and other

XX haga repeat unit products (see also AAR96021-23) can be obtd. from

XX CC transformed host cells and used as vaccines to protect humans or

XX CC animals against periodontal disease. Expression in Salmonella

XX cells allows prodn. of live vaccine. Harep4-4 can also be used

XX CC to detect the presence of anti-P. gingivalis antibodies and to

XX CC raise monoclonal antibodies for diagnostic appln.

XX Sequence 439 AA:

SO Query Match 97.4%: Score 711: DB 17: Length 439:

Best Local Similarity 97.0%: Pred. No. 3.7e-65:

Matches 130: Conservative 3: Mismatches 1: Indels 0: Gaps 0:

QY 1 ADFTFEESSTHGEAPAEWTTIDADGEGWCLSSGQDLMTLAHGTNVVSSFSWNGMA 60

```

Db      192  ADFETFEFSTHGEPAEWTTIDADGCGWCLSSGQLDMLTAHGCTNVVASFSGMGA 251
OY      61  LNPDNVYLISKDVYTGATKVKYYPVNDGFPDGHYAVMISKGTGNAGDFTVVEETPNGINK 120
        252  LNPDNVYLISKDVYTGATKVKYYPVNDGFPDGHYAVMISKGTGNAGDFTVVEETPNGINK 311
Db      121  GGARFGLSTEANGA 134
        312  GGARFGLSTEADGA 325
OY
Db
RESULT 13
AAM69492 standard; Protein: 439 AA.
AAM69492:
AC      AAM69492:
XX
XX      22-DEC-1998 (first entry)
XX
XX      Haemagglutinin protein ha9A, HAREP4.
XX
XX      Haemagglutinin protein; periodontal disease; vaccine; ha9A.
XX
XX      Porphyromonas gingivalis.
XX
XX      US5824791-A.
XX
XX      20-OCT-1998.
XX
XX      11-DEC-1995; 95US-0570311.
XX
XX      11-DEC-1995; 95US-0570311.
XX
XX      08-SEP-1998; 88US-0241640.
XX
XX      25-JAN-1991; 91US-0647119.
XX
XX      09-DEC-1994; 94US-0353485.
XX
XX      (UABR-) UAB RES FOUND.
XX
XX      (UFL) UNIV FLORIDA.
XX
XX      Han N, Lantz M, Lepine G, Patel JM, Proguliske-Fox A;
XX
XX      Tummasorn S;
XX
XX      WPI: 1998-582627/49.
XX
XX      N-PSDB; AAV58879.
XX
XX      Isolated Porphyromonas gingivalis genes - encoding haemagglutinin
XX
XX      and/or protease poly:peptide(s))
XX
XX      Claim 1; Column 139-144; 101pp; English.
XX
XX      This sequence is encoded by a Porphyromonas gingivalis gene of the
XX
XX      invention. This sequence represents the ha9A haemagglutinin protein. The
XX
XX      polypeptides are used to produce antibodies to organisms associated with
XX
XX      periodontal disease. The antibodies are also used in purification and
XX
XX      identification procedures. The genes and polypeptides are used as
XX
XX      vaccines against periodontal disease.
XX
XX      Sequence 439 AA:
SQ
Query Match      97.4%; Score 711; DB 19; Length 439;
Best Local Similarity 97.0%; Pred. No. 3,7e-65;
Matches 130; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
OY      1  ADFETFEFSTHGEPAEWTTIDADGCGWCLSSGQLDMLTAHGCTNVVASFSGMGA 60
        192  ADFETFEFSTHGEPAEWTTIDADGCGWCLSSGQLDMLTAHGCTNVVASFSGMGA 251
Db      61  LNPDNVYLISKDVYTGATKVKYYPVNDGFPDGHYAVMISKGTGNAGDFTVVEETPNGINK 120
        252  LNPDNVYLISKDVYTGATKVKYYPVNDGFPDGHYAVMISKGTGNAGDFTVVEETPNGINK 311
OY      121  GGARFGLSTEANGA 134
        312  GGARFGLSTEADGA 134
OY

```

```

Db      312  GGARFGLSTEADGA 325
OY
Db
RESULT 14
AAR96021 standard; Protein: 450 AA.
AAR96021:
AC      AAR96021:
XX
XX      04-SEP-1996 (first entry)
XX
XX      P. gingivalis haemagglutinin ha9A HAREP1 product.
XX
XX      Haemagglutinin; ha9A; periodontal disease; vaccine; antibody;
XX
XX      HAREP1.
XX
XX      Porphyromonas gingivalis strain 381.
XX
XX      WO9617936-A2.
XX
XX      13-JUN-1996.
XX
XX      11-DEC-1995; 95WO-US16108.
XX
XX      09-DEC-1994; 94US-0353485.
XX
XX      (UABR-) UAB RES FOUND.
XX
XX      (UFL) UNIV FLORIDA.
XX
XX      Han N, Lantz M, Lepine G, Patel JM, Proguliske-Fox A;
XX
XX      Tummasorn S;
XX
XX      WPI: 1996-287181/29.
XX
XX      N-PSDB; AAT30645.
XX
XX      Porphyromonas gingivalis genes and proteins - used in the detection
XX
XX      and vaccination against periodontal disease
XX
XX      Claim 4; Page 103-104; 153pp; English.
XX
XX      HAREP1 (AAR96021) is the product of the HAREP1 repeat unit (AAT30645)
XX
XX      of the ha9A gene (AAT30654) of P. gingivalis 318. It forms part
XX
XX      of haemagglutinin ha9A (see also AAR96030). HAREP1 and other
XX
XX      ha9A repeat unit products (see also AAR96022-24) can be obtained from
XX
XX      transformed host cells and used as vaccines to protect humans or
XX
XX      animals against periodontal disease. Expression in Salmonella
XX
XX      cells allows production of live vaccine. HAREP1-4 can also be used
XX
XX      to detect the presence of anti-P. gingivalis antibodies and to
XX
XX      raise monoclonal antibodies for diagnostic application.
XX
XX      Sequence 450 AA:
SQ
Query Match      97.4%; Score 711; DB 17; Length 450;
Best Local Similarity 97.0%; Pred. No. 3,9e-65;
Matches 130; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
OY      1  ADFETFEFSTHGEPAEWTTIDADGCGWCLSSGQLDMLTAHGCTNVVASFSGMGA 60
        186  ADFETFEFSTHGEPAEWTTIDADGCGWCLSSGQLDMLTAHGCTNVVASFSGMGA 245
Db      61  LNPDNVYLISKDVYTGATKVKYYPVNDGFPDGHYAVMISKGTGNAGDFTVVEETPNGINK 120
        246  LNPDNVYLISKDVYTGATKVKYYPVNDGFPDGHYAVMISKGTGNAGDFTVVEETPNGINK 305
OY      121  GGARFGLSTEANGA 134
        306  GGARFGLSTEADGA 319
OY

```

```

RESULT 15
AAM69489 standard; Protein: 450 AA.
AAM69489:
XX
XX

```


AC AAK69489;

XX 22-DEC-1998 (first entry)

XX Haemagglutinin protein hagg, HAREPI.

XX Haemagglutinin protein; periodontal disease; vaccine; hagg.

XX Porphyromonas gingivalis.

XX US5824791-A.

XX 20-OCT-1998.

XX 11-DEC-1995; 95US-0570311.

XX 11-DEC-1995; 95US-0570311.

XX 08-SEP-1988; 88US-0241640.

XX 23-JAN-1991; 91US-0647119.

XX 09-DEC-1994; 94US-0353485.

XX (UABR-) UAB RES FOUND.

XX (UYFL) UNIV FLORIDA.

XX Han N, Lantz M, Lepline G, Patcl JM, Proguiske-Fox A;

XX Tumwasorn S.

XX WPI; 1998-58267/49.

XX N-PSDB; AAV58876.

XX Isolated Porphyromonas gingivalis genes - encoding haemagglutinin

XX and/or protease poly(peptide(s))

XX Claim 1; Column 121-126; 101pp; English.

XX This sequence is encoded by a Porphyromonas gingivalis gene of the

XX invention. This sequence represents the hagg haemagglutinin protein. The

XX polypeptides are used to produce antibodies to organisms associated with

XX periodontal disease. The antibodies are also used in purification and

XX identification procedures. The genes and polypeptides are used as

XX vaccines against periodontal disease.

XX Sequence 450 AA:

SQ

Query Match 97.4%; Score 711; DB 19; Length 450;

Best Local Similarity 97.0%; Pred. No. 3.9e-65;

Matches 130; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 ADFTTFESSTGGEAPAEWTTIDADGEGWICLSSGQIDWLTAGGTNNVSSFSMNGMA 60

Db 186 ADFTTFESSTGGEAPAEWTTIDADGEGWICLSSGQIDWLTAGGTNNVSSFSMNGMA 245

QY 61 LNPDMNLISKDVYTGATKVVYYPVNDGFFGDHYAVMISKGTNAGDFTVVEETPNKINK 120

Db 246 LNPDMNLISKDVYTGATKVVYYPVNDGFFGDHYAVMISKGTNAGDFTVVEETPNKINK 305

QY 121 GGAFFGLSTLANGA 134

Db 306 GGAFFGLSTLANGA 319

Search completed: June 26, 2003, 00:51:42

Job time: 47.2961 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 26, 2003, 00:50:09 ; Search time 19.7059 Seconds
(without alignments)
200.076 Million cell updates/sec

Title: US-09-980-370-6
Perfect score: 730
Sequence: 1 ADPFTFESSHGEAPAEWT.....PMGINKGARFGLSTSEANCA 134

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PTC05_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfillsl.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	716	98.1	1687	2	US-08-570-311-29	Sequence 29, App1
2	716	98.1	1704	3	US-08-336-308A-10	Sequence 10, App1
3	716	98.1	1704	3	US-08-822-324-6	Sequence 6, App1
4	716	98.1	1704	4	US-09-490-931-10	Sequence 10, App1
5	715	97.9	1732	2	US-08-570-311-10	Sequence 10, App1
6	715	97.9	1732	2	US-08-353-485-10	Sequence 10, App1
7	711	97.4	439	2	US-08-570-311-22	Sequence 22, App1
8	711	97.4	450	2	US-08-570-311-16	Sequence 16, App1
9	711	97.4	2628	2	US-08-570-311-14	Sequence 14, App1
10	709	97.1	1087	2	US-08-570-311-8	Sequence 8, App1
11	709	97.1	1087	2	US-08-353-485-8	Sequence 8, App1
12	709	97.1	1358	2	US-08-570-311-27	Sequence 27, App1
13	704	96.4	456	2	US-08-570-311-18	Sequence 18, App1
14	704	96.4	456	2	US-08-570-311-20	Sequence 20, App1
15	176.5	24.2	497	2	US-08-570-311-2	Sequence 2, App1
16	176.5	24.2	497	2	US-08-353-485-2	Sequence 2, App1
17	83.5	11.4	316	3	US-08-682-643-4	Sequence 4, App1
18	79.5	10.9	316	1	US-08-038-932B-1	Sequence 1, App1
19	79.5	10.9	316	1	US-08-656-349-1	Sequence 1, App1
20	79.5	10.9	316	4	US-09-104-623A-5	Sequence 5, App1
21	79.5	10.9	316	4	US-09-019-532-5	Sequence 5, App1
22	78	10.7	785	3	US-09-265-108-2	Sequence 2, App1
23	78	10.7	785	4	US-09-479-264-2	Sequence 2, App1
24	73.5	10.1	319	3	US-08-682-643-3	Sequence 3, App1
25	72.5	9.9	435	4	US-09-072-917A-9	Sequence 9, App1
26	72.5	9.9	265	3	US-08-483-857-8	Sequence 8, App1
27	72.5	9.9	617	4	US-09-188-930-303	Sequence 303, App1

28	71	9.7	236	3	US-08-886-269-2	Sequence 2, App1
29	71	9.7	236	3	US-09-167-647-2	Sequence 2, App1
30	71	9.7	758	2	US-08-560-398-2	Sequence 10, App1
31	69.5	9.5	2123	4	US-08-968-685A-10	Sequence 33, App1
32	69	9.5	271	3	US-08-968-563-33	Sequence 33, App1
33	69	9.5	271	3	US-08-969-683A-33	Sequence 7, App1
34	69	9.5	271	4	US-09-297-928-14	Sequence 14, App1
35	68	9.3	250	3	US-08-968-563-17	Sequence 17, App1
36	68	9.3	250	4	US-08-969-683A-17	Sequence 13, App1
37	68	9.3	250	4	US-09-297-928-13	Sequence 13, App1
38	68	9.3	288	2	US-08-875-062-1	Sequence 3, App1
39	68	9.3	338	1	US-08-606-888A-3	Sequence 7, App1
40	68	9.3	531	2	US-07-862-588B-7	Sequence 33, App1
41	67	9.2	2353	4	US-09-377-155-33	Sequence 33, App1
42	67	9.2	2353	4	US-08-913-942-4	Sequence 4, App1
43	67	9.2	2353	4	US-09-669-974-33	Sequence 33, App1
44	67	9.2	2354	4	US-09-268-347-47	Sequence 47, App1
45	67	9.2	2411	4	US-09-268-347-36	Sequence 36, App1

ALIGNMENTS

RESULT 1
US-08-570-311-29
Sequence 29, Application US/08570311
Patent No. 5824791
GENERAL INFORMATION:
APPLICANT: Proguiske-Fox, Ann
APPLICANT: Tumwasorn, Somying
APPLICANT: Lepine, Guylaine
APPLICANT: Han, Naiming
APPLICANT: Lantz, Marilyn
APPLICANT: Patti, Joseph
TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
TITLE OF INVENTION: and probes for the Detection of Periodontal Disease
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Ted W. Whitlock
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION NUMBER: US/08/570,311
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/353,485
FILING DATE: 09-DEC-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/647,119
FILING DATE: 25-JAN-1991
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/241,640
FILING DATE: 08-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UF15_C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:

LENGTH: 1687 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-570-311-29

Query Match 98.1%; Score 716; DB 2; Length 1687;
Best Local Similarity 97.8%; Pred. No. 3.8e-67;

Matches 131; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ADETFESSSTHGEAPAEWTTIDADDGEGWCLSSGQDMLTAHGTNNVSSFSNNGMA 60
DB 1120 ADETFESSSTHGEAPAEWTTIDADDGEGWCLSSGQDMLTAHGTNNVSSFSNNGMA 1179
QY 61 LNPDMYLSKDYTGATKVKYYPVNDGFPGDHYAVAMISKTGTNAGDFTVVEETPNKINK 120
DB 1180 LNPDMYLSKDYTGATKVKYYPVNDGFPGDHYAVAMISKTGTNAGDFTVVEETPNKINK 1239
QY 121 GGARFGLSTEANGA 134
DB 1240 GGARFGLSTEANGA 1253

RESULT 2

US-08-336-308A-10
Sequence 10, Application US/08336308A
Patent No. 6017532

GENERAL INFORMATION:

APPLICANT: Travis, James
APPLICANT: Potempa, Jan S.
APPLICANT: Barr, Philip J.
APPLICANT: Pavloff, Nadine
TITLE OF INVENTION: Porphyromonas gingivalis
TITLE OF INVENTION: Arginine-specific Proteinase Coding Sequences
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: US
ZIP: 80303

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/336,308A
FILING DATE: 08-NOV-1994

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/119,361
FILING DATE: 10-SEP-1993

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/265,441
FILING DATE: 24-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Felder, Donna M.
REGISTRATION NUMBER: 33,878
REFERENCE/DOCKET NUMBER: 21-93C

TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 499-8080
TELEFAX: (303) 499-8089
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1704 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-336-308A-10

Query Match 98.1%; Score 716; DB 3; Length 1704;

Best Local Similarity 97.8%; Pred. No. 3.8e-67;
Matches 131; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ADETFESSSTHGEAPAEWTTIDADDGEGWCLSSGQDMLTAHGTNNVSSFSNNGMA 60
DB 1137 ADETFESSSTHGEAPAEWTTIDADDGEGWCLSSGQDMLTAHGTNNVSSFSNNGMA 1196
QY 61 LNPDMYLSKDYTGATKVKYYPVNDGFPGDHYAVAMISKTGTNAGDFTVVEETPNKINK 120
DB 1197 LNPDMYLSKDYTGATKVKYYPVNDGFPGDHYAVAMISKTGTNAGDFTVVEETPNKINK 1256
QY 121 GGARFGLSTEANGA 134
DB 1257 GGARFGLSTEANGA 1270

RESULT 3

US-08-822-324-6
Sequence 6, Application US/08822324
Patent No. 6129917

GENERAL INFORMATION:

APPLICANT: Potempa, Jan S.
APPLICANT: Travis, James
APPLICANT: Genco, Caroline A.
TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS COMPRISING
TITLE OF INVENTION: PORPHYROMONAS GINGIVALIS PROTEINS AND/OR PEPTIDES AND
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: CO
COUNTRY: US
ZIP: 80303

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/822,324
FILING DATE: 21-MAR-1997
CLASSIFICATION: 536

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/013,945
FILING DATE: 22-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Felder, Donna M.
REGISTRATION NUMBER: 33,878
REFERENCE/DOCKET NUMBER: 103-95 WO

TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 488-8080
TELEFAX: (303) 499-8089
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1704 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-822-324-6

Query Match 98.1%; Score 716; DB 3; Length 1704;
Best Local Similarity 97.8%; Pred. No. 3.8e-67;
Matches 131; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ADETFESSSTHGEAPAEWTTIDADDGEGWCLSSGQDMLTAHGTNNVSSFSNNGMA 60
DB 1137 ADETFESSSTHGEAPAEWTTIDADDGEGWCLSSGQDMLTAHGTNNVSSFSNNGMA 1196
QY 61 LNPDMYLSKDYTGATKVKYYPVNDGFPGDHYAVAMISKTGTNAGDFTVVEETPNKINK 120
DB 1197 LNPDMYLSKDYTGATKVKYYPVNDGFPGDHYAVAMISKTGTNAGDFTVVEETPNKINK 1256

OY 121 GGAFFGLSTEANGA 134
DB 1257 GGAFFGLSTEANGA 1270

RESULT 4

US-09-931-10
Sequence 10, Application US/09490931
Patent No. 6274718

GENERAL INFORMATION:

APPLICANT: Treavis, James
APPLICANT: Potempa, Jan S.
APPLICANT: Barr, Phillip J.
APPLICANT: Pavloff, Nadine
TITLE OF INVENTION: Porphyromonas gingivalis
TITLE OF INVENTION: Arginine-specific Proteinase Coding Sequences
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: US
ZIP: 80303

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/490,931
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/336,308
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/265,441
FILING DATE: 24-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Ferber, Donna M.
REGISTRATION NUMBER: 33,878
REFERENCE/DOCKET NUMBER: 21-93C
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 499-8080
TELEFAX: (303) 499-8080
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1704 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-931-10

US-09-931-10

Query Match 98.1%; Score 716; DB 4; Length 1704;
Best Local Similarity 97.8%; Pred. No. 3.8e-67;
Matches 131; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 ADFTEFEESSTHGEPAEWTTTIDAGDGGWCLSSGOLDMLTAHGTNNVSSSSNGMA 60
DB 1137 ADFTEFEESSTHGEPAEWTTTIDAGDGGWCLSSGOLDMLTAHGTNNVSSSSNGMA 1196
OY 61 LNPBNVLYLSKDVATGATKVKYYPVNDGFGPDHVAVMISKTGTNAGDFTVVEETPNGLNK 120
DB 1197 LNPBNVLYLSKDVATGATKVKYYPVNDGFGPDHVAVMISKTGTNAGDFTVVEETPNGLNK 1256
OY 121 GGAFFGLSTEANGA 134
DB 1257 GGAFFGLSTEANGA 1270

RESULT 5

US-08-570-311-10
Sequence 10, Application US/08570311
Patent No. 5824791

GENERAL INFORMATION:

APPLICANT: Progulsk-Fox, Ann
APPLICANT: Tumwasorn, Somying
APPLICANT: Lepine, Guylaine
APPLICANT: Han, Naiming
APPLICANT: Lantz, Marilyn
APPLICANT: Patti, Joseph
TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ted W. Whitlock
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/570,311
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/353,485
FILING DATE: 09-DEC-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/647,119
FILING DATE: 25-JAN-1991
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/241,640
FILING DATE: 08-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: 0F15.C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1732 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-570-311-10

US-08-570-311-10

Query Match 97.9%; Score 715; DB 2; Length 1732;
Best Local Similarity 97.8%; Pred. No. 5e-67;
Matches 131; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 ADFTEFEESSTHGEPAEWTTTIDAGDGGWCLSSGOLDMLTAHGTNNVSSSSNGMA 60
DB 1157 ADFTEFEESSTHGEPAEWTTTIDAGDGGWCLSSGOLDMLTAHGTNNVSSSSNGMA 1216
OY 61 LNPBNVLYLSKDVATGATKVKYYPVNDGFGPDHVAVMISKTGTNAGDFTVVEETPNGLNK 120
DB 1217 LNPBNVLYLSKDVATGATKVKYYPVNDGFGPDHVAVMISKTGTNAGDFTVVEETPNGLNK 1276
OY 121 GGAFFGLSTEANGA 134
DB 1277 GGAFFGLSTEANGA 1290

RESULT 6

US-08-353-485-10

Sequence 10, Application US/08353485

Patent No. 5830710

GENERAL INFORMATION:

APPLICANT: Proguliske-Fox, Ann

APPLICANT: Tumwasorn, Somying

APPLICANT: Lepine, Guylaine

APPLICANT: Han, Naiming

APPLICANT: Lantz, Marilyn

APPLICANT: Patti, Joseph

TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes

TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Ted W. Whitlock

STREET: 2421 N.W. 41st Street, Suite A-1

CITY: Gainesville

STATE: FL

COUNTRY: USA

ZIP: 32606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/353,485

FILING DATE: 09-DEC-1994

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/647,119

FILING DATE: 25-JAN-1991

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/241,640

FILING DATE: 08-SEP-1988

ATTORNEY/AGENT INFORMATION:

NAME: Whitlock, Ted W.

REGISTRATION NUMBER: 36,965

REFERENCE/DOCKET NUMBER: UF15.C2

TELECOMMUNICATION INFORMATION:

TELEPHONE: (904) 375-8100

TELEFAX: (904) 372-5800

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 1732 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-353-485-10

Query Match 97.9%; Score 715; DB 2; Length 1732;

Best Local Similarity 97.8%; Pred. No. 5e-67;

Matches 131; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ADETFEESSTHGEAPAEWTTTIDADGDEGMLCLSSGOLDMLTAHAGTNNVSSFSWNGMA 60

DB 1157 ADETFEESSTHGEAPAEWTTTIDADGDEGMLCLSSGOLDMLTAHAGTNNVSSFSWNGMA 1216

QY 61 LNPDMYLLSKDVTGATKRYKYYYPVNDGFPGDHYAVMISKTGTNAGDFTVVFETPNKINK 120

DB 1217 LNPDMYLLSKDVTGATKRYKYYYPVNDGFPGDHYAVMISKTGTNAGDFTVVFETPNKINK 1276

QY 121 GGAFFGLSTEANGA 134

DB 1277 GGAFFGLSTEANGA 1290

RESULT 7

US-08-570-311-22

Sequence 22, Application US/08570311

Patent No. 5824791

GENERAL INFORMATION:

APPLICANT: Proguliske-Fox, Ann

APPLICANT: Tumwasorn, Somying

APPLICANT: Lepine, Guylaine

APPLICANT: Han, Naiming

APPLICANT: Lantz, Marilyn

APPLICANT: Patti, Joseph

TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes

TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease

NUMBER OF SEQUENCES: 29

CORRESPONDENCE ADDRESS:

ADDRESSEE: Ted W. Whitlock

STREET: 2421 N.W. 41st Street, Suite A-1

CITY: Gainesville

STATE: FL

COUNTRY: USA

ZIP: 32606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/570,311

FILING DATE:

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/353,485

FILING DATE: 09-DEC-1994

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/647,119

FILING DATE: 25-JAN-1991

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/241,640

FILING DATE: 08-SEP-1988

ATTORNEY/AGENT INFORMATION:

NAME: Whitlock, Ted W.

REGISTRATION NUMBER: 36,965

REFERENCE/DOCKET NUMBER: UF15.C3

TELECOMMUNICATION INFORMATION:

TELEPHONE: (904) 375-8100

TELEFAX: (904) 372-5800

INFORMATION FOR SEQ ID NO: 22:

SEQUENCE CHARACTERISTICS:

LENGTH: 439 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-570-311-22

Query Match 97.4%; Score 711; DB 2; Length 439;

Best Local Similarity 97.0%; Pred. No. 2e-67;

Matches 130; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 ADETFEESSTHGEAPAEWTTTIDADGDEGMLCLSSGOLDMLTAHAGTNNVSSFSWNGMA 60

DB 192 ADETFEESSTHGEAPAEWTTTIDADGDEGMLCLSSGOLDMLTAHAGTNNVSSFSWNGMA 251

QY 61 LNPDMYLLSKDVTGATKRYKYYYPVNDGFPGDHYAVMISKTGTNAGDFTVVFETPNKINK 120

DB 252 LNPDMYLLSKDVTGATKRYKYYYPVNDGFPGDHYAVMISKTGTNAGDFTVVFETPNKINK 311

QY 121 GGAFFGLSTEANGA 134

DB 312 GGAFFGLSTEANGA 325

RESULT 8

US-08-570-311-16

Sequence 16, Application US/08570311

Patent No. 5824791

GENERAL INFORMATION:

APPLICANT: Proguiske-Fox, Ann
APPLICANT: Tummasorn, Somying
APPLICANT: Lepine, Guyelaine
APPLICANT: Han, Naiming
APPLICANT: Lantz, Marilyn
APPLICANT: Patti, Joseph
TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ted W. Whitlock
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/570,311
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/353,485
FILING DATE: 09-DEC-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/647,119
FILING DATE: 25-JAN-1991
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/241,640
FILING DATE: 08-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UFI5.C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 450 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-570-311-16

Query Match 97.4%; Score 711; DB 2; Length 450;
Best Local Similarity 97.0%; Pred. No. 2e-67;
Matches 130; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 ADFTETESSHGEAPAEWTTIDADDGEGWLCSSGOLDMLTAHGTNVVVSFSWNGMA 60
DB 186 ADFTETESSHGEAPAEWTTIDADDGEGWLCSSGOLDMLTAHGTNVVVSFSWNGMA 245
QY 61 LNPNDNLISDVYGAARKVKKYYPVNDGPGDHYAVMISKTGTNAGDFTVVEETPNCINK 120
DB 246 LNPNDNLISDVYGAARKVKKYYPVNDGPGDHYAVMISKTGTNAGDFTVVEETPNCINK 305
QY 121 GGAFFGLSTEANGA 134
DB 306 GGAFFGLSTEADGA 319

RESULT 9
US-08-570-311-14
Sequence 14, Application US/08570311
Patent No. 5824791
GENERAL INFORMATION:

APPLICANT: Proguiske-Fox, Ann
APPLICANT: Tummasorn, Somying
APPLICANT: Lepine, Guyelaine
APPLICANT: Han, Naiming
APPLICANT: Lantz, Marilyn
APPLICANT: Patti, Joseph
TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ted W. Whitlock
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/570,311
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/353,485
FILING DATE: 09-DEC-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/647,119
FILING DATE: 25-JAN-1991
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/241,640
FILING DATE: 08-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UFI5.C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 2628 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-570-311-14

Query Match 97.4%; Score 711; DB 2; Length 2628;
Best Local Similarity 97.0%; Pred. No. 2.4e-66;
Matches 130; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 ADFTETESSHGEAPAEWTTIDADDGEGWLCSSGOLDMLTAHGTNVVVSFSWNGMA 60
DB 685 ADFTETESSHGEAPAEWTTIDADDGEGWLCSSGOLDMLTAHGTNVVVSFSWNGMA 744
QY 61 LNPNDNLISDVYGAARKVKKYYPVNDGPGDHYAVMISKTGTNAGDFTVVEETPNCINK 120
DB 745 LNPNDNLISDVYGAARKVKKYYPVNDGPGDHYAVMISKTGTNAGDFTVVEETPNCINK 804
QY 121 GGAFFGLSTEANGA 134
DB 805 GGAFFGLSTEADGA 818

RESULT 10
US-08-570-311-8
Sequence 8, Application US/08570311
Patent No. 5824791
GENERAL INFORMATION:

APPLICANT: Proguiske-Fox, Ann
APPLICANT: Tumwasorn, Somying
APPLICANT: Lepine, Guyalaine
APPLICANT: Han, Naiming
APPLICANT: Lantz, Marilyn
APPLICANT: Patti, Joseph
TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ted W. Whitlock
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/570,311
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/353,485
FILING DATE: 09-DEC-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/647,119
FILING DATE: 25-JAN-1991
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/241,640
FILING DATE: 08-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UFI5.C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEO ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1087 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-570-311-8

Query Match 97.1%; Score 709; DB 2; Length 1087;
Best Local Similarity 97.0%; Pred. No. 1.1e-66;
Matches 130; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ADFTTFESSSTHGEAPAEWTTIDADGDEGWLCLSSGQLDWLTAHGTVNVSSFSWNGMA 60
DB 520 ADFTTFESSSTHGEAPAEWTTIDADGDEGWLCLSSGQLDWLTAHGTVNVSSFSWNGMA 579
QY 61 LNPDNVLLSKDYTGATKTKKYYYPVNDGPGDHYAVMISKTGTNAGDFTVVFETPNCINK 120
DB 580 LNPDNVLLSKDYTGATKTKKYYYPVNDGPGDHYAVMISKTGTNAGDFTVVFETPNCINK 639
QY 121 GGARFGLSTEANCA 134
DB 640 GGARFGLSTEANCA 653

RESULT 11
US-08-353-485-8
Sequence 8, Application US/08353485
Patent No. 5830710
GENERAL INFORMATION:

APPLICANT: Proguiske-Fox, Ann
APPLICANT: Tumwasorn, Somying
APPLICANT: Lepine, Guyalaine
APPLICANT: Han, Naiming
APPLICANT: Lantz, Marilyn
APPLICANT: Patti, Joseph
TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ted W. Whitlock
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/353,485
FILING DATE: 09-DEC-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/647,119
FILING DATE: 25-JAN-1991
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/241,640
FILING DATE: 08-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UFI5.C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEO ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1087 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-353-485-8

Query Match 97.1%; Score 709; DB 2; Length 1087;
Best Local Similarity 97.0%; Pred. No. 1.1e-66;
Matches 130; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ADFTTFESSSTHGEAPAEWTTIDADGDEGWLCLSSGQLDWLTAHGTVNVSSFSWNGMA 60
DB 520 ADFTTFESSSTHGEAPAEWTTIDADGDEGWLCLSSGQLDWLTAHGTVNVSSFSWNGMA 579
QY 61 LNPDNVLLSKDYTGATKTKKYYYPVNDGPGDHYAVMISKTGTNAGDFTVVFETPNCINK 120
DB 580 LNPDNVLLSKDYTGATKTKKYYYPVNDGPGDHYAVMISKTGTNAGDFTVVFETPNCINK 639
QY 121 GGARFGLSTEANCA 134
DB 640 GGARFGLSTEANCA 653

RESULT 12
US-08-570-311-27
Sequence 27, Application US/08570311
Patent No. 5824791
GENERAL INFORMATION:
APPLICANT: Proguiske-Fox, Ann
APPLICANT: Tumwasorn, Somying
APPLICANT: Lepine, Guyalaine
APPLICANT: Han, Naiming

APPLICANT: Lantz, Marilyn
APPLICANT: Patti, Joseph
TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ted W. Whitlock
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/570,311
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/353,485
FILING DATE: 09-DEC-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/647,119
FILING DATE: 25-JAN-1991
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/241,640
FILING DATE: 08-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UFI5.C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 1358 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-570-311-27

Query Match 97.1%; Score 709; DB 2; Length 1358;
Best Local Similarity 97.0%; Pred. No. 1.5e-66;
Matches 130; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ADFTETFESSSTHGEAPAEWTTTIDADGDEGWLCLSSGQLDMLTAHGTVNVSSFSWNGMA 60
DB 791 ADFTETFESSSTHGEAPAEWTTTIDADGDEGWLCLSSGQLDMLTAHGTVNVSSFSWNGMA 850
QY 61 LNPDPNYLSKDVYTGATKTKYKYYVPVNDGFPDGHYAVMISKGTNAGDTTVVFEETPNGINK 120
DB 851 LNPDPNYLSKDVYTGATKTKYKYYVPVNDGFPDGHYAVMISKGTNAGDTTVVFEETPNGINK 910
QY 121 GGARFGLSTEANGA 134
DB 911 GGARFGLSTEANGA 924

RESULT 13
US-08-570-311-18
Sequence 18, Application US/08570311
Patent No. 5824791
GENERAL INFORMATION:
APPLICANT: Proguiske-Fox, Ann
APPLICANT: Tumwasorn, Somying
APPLICANT: Lepine, Guylaine
APPLICANT: Han, Naimeing

APPLICANT: Lantz, Marilyn
APPLICANT: Patti, Joseph
TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ted W. Whitlock
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/570,311
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/353,485
FILING DATE: 09-DEC-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/647,119
FILING DATE: 25-JAN-1991
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/241,640
FILING DATE: 08-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UFI5.C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 456 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-570-311-18

Query Match 96.4%; Score 704; DB 2; Length 456;
Best Local Similarity 96.3%; Pred. No. 1.1e-66;
Matches 129; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 ADFTETFESSSTHGEAPAEWTTTIDADGDEGWLCLSSGQLDMLTAHGTVNVSSFSWNGMA 60
DB 192 ADFTETFESSSTHGEAPAEWTTTIDADGDEGWLCLSSGQLDMLTAHGTVNVSSFSWNGMA 251
QY 61 LNPDPNYLSKDVYTGATKTKYKYYVPVNDGFPDGHYAVMISKGTNAGDTTVVFEETPNGINK 120
DB 252 LNPDPNYLSKDVYTGATKTKYKYYVPVNDGFPDGHYAVMISKGTNAGDTTVVFEETPNGINK 311
QY 121 GGARFGLSTEANGA 134
DB 312 GGARFGLSTEANGA 325

RESULT 14
US-08-570-311-20
Sequence 20, Application US/08570311
Patent No. 5824791
GENERAL INFORMATION:
APPLICANT: Proguiske-Fox, Ann
APPLICANT: Tumwasorn, Somying
APPLICANT: Lepine, Guylaine
APPLICANT: Han, Naimeing

```
APPLICANT: Lantz, Marilyn
APPLICANT: Patti, Joseph
TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESS: Ted W. Whitlock
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/570,311
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/353,485
FILING DATE: 09-DEC-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/647,119
FILING DATE: 25-JAN-1991
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/241,640
FILING DATE: 08-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UF15.C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 456 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-570-311-20

Query Match          96.4%; Score 704; DB 2; Length 456;
Best Local Similarity 96.3%; Pred. No. 1.1e-66;
Matches 129; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 ADFTETFEESSTHGEAPAEWTTTIDADGEGWMLCLSSGOLDMLTAHGCTNVVSSFSNMGMA 60
    |:::| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 192 ADFTETFEESSTHGEAPAEWTTTIDADGEGWMLCLSSGOLDMLTAHGCTNVVSSFSNMGMA 251
    |:::| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 61 LNPDMVLLSKDYTGATKVKYKYYVNDGFPGDHVAVMISKTGTNAGDPTVVEETPNKINK 120
    |:::| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 252 LNPDMVLLSKDYTGATKVKYKYYVNDGFPGDHVAVMISKTGTNAGDPTVVEETPNKINK 311
    |:::| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 121 GGARFGLSTEANGA 134
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 312 GGARFGLSTEANGA 325
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 15
US-08-570-311-2
: Sequence 2, Application US/08570311
: Patent No. 5824791
: GENERAL INFORMATION:
: APPLICANT: Proguliske-Fox, Ann
: APPLICANT: Tumwasorn, Somying
: APPLICANT: Lepline, Guyalaine
: APPLICANT: Han, Nalming
```

```
APPLICANT: Lantz, Marilyn
APPLICANT: Patti, Joseph
TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESS: Ted W. Whitlock
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/570,311
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/353,485
FILING DATE: 09-DEC-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/647,119
FILING DATE: 25-JAN-1991
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/241,640
FILING DATE: 08-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UF15.C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 497 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-570-311-2

Query Match          24.2%; Score 176.5; DB 2; Length 497;
Best Local Similarity 38.1%; Pred. No. 1.2e-10;
Matches 45; Conservative 16; Mismatches 44; Indels 13; Gaps 5;
```

```
QY 5 ETEFESSTHGEAPAEWTTTIDADGEGWMLCLSSGOLDMLTAHGCTNVVSSFSNMGMA 59
    |:::| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 235 ESFDRQT--LPNGWTMIDADGDGHNW--LSTINVTNTATHTGDDGAMFESKSWTASGAKI 289
    |:::| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 60 ALNPDNVLISKDYTGATKVKYKYYVNDGFP--GDHVAVMISKTGTNAGDPTV-VEEET 114
    |:::| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 290 DLSPDNVLTPKVTVPENCKLSYWSQYPTNEHGVLPSTTGNEANETTKLLEET 347
    |:::| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

Search completed: June 26, 2003, 00:53:58
Job time : 19.7059 secs

Thu Jun 26 11:58:15 2003

us-09-980-370-6.rapb

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 26, 2003, 00:50:39 ; Search time 28.3765 Seconds

(without alignments)
510.977 Million cell updates/sec

Title: US-09-980-370-6

Sequence: 1 ADPTEFSSHGEAFAEWTTIDAGDGCDEMLCTSSGQIDWLTAGCTNVSSFS--WNCHAA 60

Scoring table:

BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 417779 seqs, 108206813 residues

Total number of hits satisfying chosen parameters: 417779

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/pdata1/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/pdata1/pubpaa/US06_NEW_PUB.pep.*
- 3: /cgn2_6/pdata1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/pdata1/pubpaa/US07_NEW_PUB.pep.*
- 5: /cgn2_6/pdata1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/pdata1/pubpaa/US07_NEW_PUB.pep.*
- 7: /cgn2_6/pdata1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/pdata1/pubpaa/US08_NEW_PUB.pep.*
- 9: /cgn2_6/pdata1/pubpaa/US08_NEW_PUB.pep.*
- 10: /cgn2_6/pdata1/pubpaa/US08_NEW_PUB.pep.*
- 11: /cgn2_6/pdata1/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/pdata1/pubpaa/US10_NEW_PUB.pep.*
- 13: /cgn2_6/pdata1/pubpaa/US10_NEW_PUB.pep.*
- 14: /cgn2_6/pdata1/pubpaa/US10_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	72.5	9.9	617	9 US-09-866-050A-303	Sequence 303, App
2	70	9.6	377	10 US-09-784-877-2	Sequence 2, Appl
3	70	9.6	377	10 US-09-817-913-15	Sequence 15, Appl
4	70	9.6	377	10 US-09-817-913-15	Sequence 15, Appl
5	69	9.5	371	9 US-09-308-207-33	Sequence 33, Appl
6	68	9.3	250	9 US-09-308-207-17	Sequence 17, Appl
7	68	9.3	693	10 US-09-815-242-13233	Sequence 13233, A
8	67.5	9.2	506	9 US-10-125-692-18	Sequence 18, Appl
9	67	9.2	256	9 US-09-880-748-1745	Sequence 1745, Ap
10	67	9.2	2353	10 US-09-797-862-33	Sequence 33, Appl
11	66	9.0	221	9 US-10-112-527-3	Sequence 3, Appl
12	66	9.0	236	9 US-10-112-527-2	Sequence 2, Appl
13	66	9.0	251	9 US-10-112-527-1	Sequence 1, Appl
14	66	9.0	283	9 US-10-141-531-26	Sequence 26, Appl
15	66	9.0	334	9 US-10-141-531-68	Sequence 68, Appl
16	66	9.0	334	9 US-10-141-531-81	Sequence 81, Appl
17	66	9.0	334	9 US-10-141-531-90	Sequence 90, Appl
18	66	9.0	334	9 US-10-141-531-99	Sequence 99, Appl
19	66	9.0	334	9 US-10-141-531-108	Sequence 108, App

20	66	9.0	334	9 US-10-141-531-117	Sequence 117, App
21	66	9.0	334	9 US-10-141-531-126	Sequence 126, App
22	66	9.0	334	9 US-10-141-531-135	Sequence 135, App
23	66	9.0	334	9 US-10-141-531-144	Sequence 144, App
24	66	9.0	334	9 US-10-141-531-153	Sequence 153, App
25	66	9.0	334	9 US-10-141-531-162	Sequence 162, App
26	66	9.0	334	9 US-10-141-531-171	Sequence 171, App
27	66	9.0	334	9 US-10-141-531-180	Sequence 180, App
28	66	9.0	334	9 US-10-141-531-189	Sequence 189, App
29	66	9.0	334	9 US-10-141-531-198	Sequence 198, App
30	66	9.0	575	9 US-09-738-626-4263	Sequence 4263, Ap
31	66	9.0	631	9 US-10-124-880-16	Sequence 16, Appl
32	66	9.0	847	9 US-10-112-527-4	Sequence 4, Appl
33	65.5	9.0	410	9 US-10-125-692-31	Sequence 31, Appl
34	65.5	9.0	410	9 US-09-916-494A-10	Sequence 10, Appl
35	65.5	9.0	441	9 US-09-738-626-6513	Sequence 6513, Ap
36	65	8.9	263	9 US-09-738-626-5554	Sequence 5554, Ap
37	65	8.9	304	9 US-09-984-245-200	Sequence 200, App
38	65	8.9	315	9 US-09-984-245-200	Sequence 200, App
39	65	8.9	315	9 US-09-984-245-200	Sequence 200, App
40	65	8.9	315	9 US-09-984-245-200	Sequence 200, App
41	65	8.9	315	9 US-09-984-245-200	Sequence 200, App
42	65	8.9	315	9 US-09-984-245-200	Sequence 200, App
43	65	8.9	315	9 US-09-984-245-200	Sequence 200, App
44	65	8.9	315	9 US-09-984-245-200	Sequence 200, App
45	65	8.9	315	9 US-09-984-245-200	Sequence 200, App

ALIGNMENTS

RESULT 1
US-09-866-050A-303
Sequence 303, Application US/09866050A
Publication No. US2003004047A1
GENERAL INFORMATION:
APPLICANT: Watson, James D.
APPLICANT: Strechian, Lorna
APPLICANT: Steeman, Matthew
APPLICANT: Orlus, Rene
APPLICANT: Murison, James G.
TITLE OF INVENTION: Compositions and Methods for Their Use
FILE REFERENCE: 11000.1011c4U
CURRENT APPLICATION NUMBER: US/09/866.050A
NUMBER OF SEQ ID NOS: 725
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 303
LENGTH: 617
TYPE: PRT
ORGANISM: Mouse
US-09-866-050A-303

Query Match 9.9%; Score 72.5; DB 9; Length 617;
Best Local Similarity 29.3%; Pred. No. 21;
Matches 22; Conservative 9; Mismatches 31; Indels 13; Gaps 3;

DB 301 FKTFSSHGEAFAEWTTIDAGDGCDEMLCTSSGQIDWLTAGCTNVSSFS--WNCHAA 60
QY 61 LNPDP-----YLYSK 70
DB 355 LSPDTPMPVYSIQ 369

RESULT 2
US-09-784-877-2
Sequence 2, Application US/09784877
Patient No. US20010012836A1
GENERAL INFORMATION:

Thu Jun 26 11:58:15 2003

us-09-980-370-6.rapb

Page 2

```

: APPLICANT: Erding Hu
: APPLICANT: Yuan Zhu
: TITLE OF INVENTION: HUMAN HISTONE DEACETYLASE GENE HD4
: FILE REFERENCE: GP-70516-C1
: CURRENT APPLICATION NUMBER: US/09/784,877
: CURRENT FILING DATE: 2001-02-16
: PRIOR APPLICATION NUMBER: 09/138,151
: PRIOR FILING DATE: 1998-08-21
: NUMBER OF SEQ ID NOS: 2
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO: 2
: LENGTH: 377
: TYPE: PRT
: ORGANISM: HOMO SAPIENS
: US-09-784-877-2

```

```

Query Match
Best Local Similarity 23.5%; Pred. No. 22;
Matches 32; Conservative 13; Mismatches 45; Indels 46; Gaps 5;

```

```

OY 42 LTAHGCTNVVSSFSNMGAL-----NPDNYLSKDYT--GATKYKY--YPVNDGFP 89
DB 177 LDLHGGVDEDAFSTSKVMTVSLHKFSFGFPFGTGDVSDVGLGKGRYTSVNVPIQDGIQ 236
OY 90 GDHY-----AVMISKTGNNAGDFTVVEETPENGINK----- 120
DB 237 DEKTYQICESVLEKEYQAFNPKAVVLDLADDTIADGDPGCSFNMTPVGISCKLKIYLOMQL 296
OY 121 -----GGARFGSLSTA 131
DB 297 ATLILGGGGYNLANTRA 312

```

RESULT 3

```

: US-09-817-913-15
: Sequence 15, Application US/09817913
: Patent No. US20020061860A1
: GENERAL INFORMATION:
: APPLICANT: LI, Zumei
: APPLICANT: Bonfils, Claire
: APPLICANT: Besteman, Jeffrey
: TITLE OF INVENTION: Inhibition of Specific Histone Deacetylase Isoforms
: FILE REFERENCE: 106101.144
: CURRENT APPLICATION NUMBER: US/09/817,913
: CURRENT FILING DATE: 2001-03-26
: PRIOR APPLICATION NUMBER: US 60/192,157
: PRIOR FILING DATE: 2000-03-24
: NUMBER OF SEQ ID NOS: 33
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO: 15
: LENGTH: 377
: TYPE: PRT
: ORGANISM: Human
: US-09-817-913-15

```

```

Query Match
Best Local Similarity 23.5%; Pred. No. 22;
Matches 32; Conservative 13; Mismatches 45; Indels 46; Gaps 5;

```

```

OY 42 LTAHGCTNVVSSFSNMGAL-----NPDNYLSKDYT--GATKYKY--YPVNDGFP 89
DB 177 LDLHGGVDEDAFSTSKVMTVSLHKFSFGFPFGTGDVSDVGLGKGRYTSVNVPIQDGIQ 236
OY 90 GDHY-----AVMISKTGNNAGDFTVVEETPENGINK----- 120
DB 237 DEKTYQICESVLEKEYQAFNPKAVVLDLADDTIADGDPGCSFNMTPVGISCKLKIYLOMQL 296
OY 121 -----GGARFGSLSTA 131
DB 297 ATLILGGGGYNLANTRA 312

```

```

: US-09-817-538-15
: Sequence 15, Application US/09817538
: Patent No. US20020137162A1
: GENERAL INFORMATION:
: APPLICANT: LI, Zumei
: APPLICANT: Bonfils, Claire
: APPLICANT: Besteman, Jeffrey
: TITLE OF INVENTION: Antisense Oligonucleotide Inhibition of Specific Histone
: FILE REFERENCE: 106101.144
: CURRENT APPLICATION NUMBER: US/09/817,538
: CURRENT FILING DATE: 2001-03-26
: PRIOR APPLICATION NUMBER: US 60/192,157
: PRIOR FILING DATE: 2000-03-24
: NUMBER OF SEQ ID NOS: 33
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO: 15
: LENGTH: 377
: TYPE: PRT
: ORGANISM: Human
: US-09-817-538-15

```

```

Query Match
Best Local Similarity 23.5%; Pred. No. 22;
Matches 32; Conservative 13; Mismatches 45; Indels 46; Gaps 5;

```

```

OY 42 LTAHGCTNVVSSFSNMGAL-----NPDNYLSKDYT--GATKYKY--YPVNDGFP 89
DB 177 LDLHGGVDEDAFSTSKVMTVSLHKFSFGFPFGTGDVSDVGLGKGRYTSVNVPIQDGIQ 236
OY 90 GDHY-----AVMISKTGNNAGDFTVVEETPENGINK----- 120
DB 237 DEKTYQICESVLEKEYQAFNPKAVVLDLADDTIADGDPGCSFNMTPVGISCKLKIYLOMQL 296
OY 121 -----GGARFGSLSTA 131
DB 297 ATLILGGGGYNLANTRA 312

```

RESULT 5

```

: US-09-308-207-33
: Sequence 33, Application US/09308207
: Publication No. US20030022323A1
: GENERAL INFORMATION:
: APPLICANT: MARIA DIAZ-TORRES ET AL.
: TITLE OF INVENTION: METHOD FOR THE RECOMBINANT
: PRODUCTION OF 1,3 PROPANEDIOL
: NUMBER OF SEQUENCES: 68
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Genencor International, Inc.
: STREET: 4 Cambridge Place
: CITY: Rochester
: STATE: NY
: COUNTRY: U.S.A
: ZIP: 14618
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: Windows
: SOFTWARE: FastSeq for Windows Version 2.0b
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/308,207
: FILING DATE: 13-May-1999
: CLASSIFICATION: <unknown>
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 60/030,601
: FILING DATE: 13-NOV-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Glaislet, Debora
: REGISTRATION NUMBER: 33,888
: REFERENCE/DOCKET NUMBER: GC 369-2

```

```
TELECOMMUNICATION INFORMATION:  
    TELEPHONE: 650-864-7620  
    TELEFAX: 650-845-6504  
  
    TELEX: <UNKNOWN>  
  
INFORMATION FOR SEQ ID NO: 33:  
SEQUENCE CHARACTERISTICS:  
    LENGTH: 271 amino acids  
    TYPE: amino acid  
    STRANDEDNESS: unknown  
    TOPOLOGY: unknown  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
ORGANISM: GPPI  
SEQUENCE DESCRIPTION: SEQ ID NO: 33:  
US-09-308-207-33
```

Query Match 9.5% Score 69 DB 9 Length 271:

	Best Local Similarity	24.3%	Pred No.	18	Mismatches	33	Conservative	22	Indels	28	Gaps	7
--	-----------------------	-------	----------	----	------------	----	--------------	----	--------	----	------	---

Oy 2 DFT-EFFSSRHGEAPAEWTTIDADGDEGMICLSSQGL--DWLTJAHGTGVVSSFSGN 57
 |||::||::|||::|||::|||::||
Db 89 DFAEEYVNKLEGELPEKYGGHSIEVPAAVLCLNALMLPREKMVAVALSQTDMAKKNFD 148
 Matches 33: Conservative 22: Mismatches 53: Indels 28: Gaps 7

Oy 58 GMAL-NPDNYILISKDY-TGAIRVKXY-----YPVDGFPGDHYAVMISKTGINAGDF 107
 :|:::||::|||::|||::|||::||
Db 149 ILIKRIPEYFITTANDVGOKRPHPEYLTKRGNGLCGPINEODP-----SKSK----- 194
 |::::|||::|||::|||::|||::||

Oy 108 TVFEETPPNGINKGA 123
 ::::|::|::|
Db 195 VVEEDAPAGIAACKA 210
 ::::|::|::|

RESULT 6
US-09-308-207-17
Sequence 17, Application US/09308207
Publication NO. US20030022323A1
GENERAL INFORMATION:
APPLICANT: MARIA DIAZ-TORRES ET AL.
TITLE OF INVENTION: METHOD FOR THE RECOMBINANT PRODUCTION OF 1,3 PROPANEDIOL
NUMBER OF SEQUENCES: 68
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genencor International, Inc.
STREET: 4 Cambridge Place
CITY: Rochester
STATE: NY
COUNTRY: U.S.A
ZIP: 14618
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: FASTSEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/308,207
FILING DATE: 13-May-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/030,601
FILING DATE: 13-Nov-1996
ATTORNEY/AGENT INFORMATION:
NAME: Glaister, Debra
REGISTRATION NUMBER: 33,888
REFERENCE/DOCKET NUMBER: GC 369-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-864-7620
TELEFAX: 650-845-6504
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 250 amino acids

```

: TYPE: amino acid
: STRANDEDNESS: unknown
: TOPOLOGY: unknown
: MOLECULE TYPE: protein
: ORIGINAL SOURCE:
: ORGANISM: GPP2
: SEQUENCE DESCRIPTION: SEQ ID NO: 17:
US-09-308-207-17

Query Match          9.3%: Score 68; DB 9; Length 250;
Best Local Similarity 24.0%: Pred. NO. 21;
Matches 35: Conservative 21; Mismatches 58; Indels 32; Gaps 8

OY 2 DF-TETFESSHGEAPAEWTTIDADGDEGMCLSSGOL---DWLTAHGTVNVSSFSFM 57
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 68 DFANEYVYVNLKLEALPEPVYGEKSTIEVPAAVNLCLANLMLPEKMAVAITSGTRDMAOKRFE 127
    : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
OY 58 GMAL-NPDNYLISKDV-TGATRKVY-----YVNDGFPDGHYAVMISKTGTNAGDF 107
    : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 128 HLGIRRPYFETIANDVKGKRPHEPYLKGRLGGLGYPIEDDP-----SKSK----- 173
    : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
OY 108 TVVEEFPNGINKGAR---FGLST 129
    ||||| | | | | | | | | | | | | | | | | | | | | | | |
Db 174 VVFEEDAPAGIAAGKACKTIIGATF 199
    ||||| | | | | | | | | | | | | | | | | | | | | | | |

RESULT 7
US-09-815-242-13233
: Sequence 13233, Application US/09815242
: Patent No. US20020061569A1
: GENERAL INFORMATION:
: APPLICANT: Haselbeck, Robert
: APPLICANT: Ohlsen, Kari L.
: APPLICANT: Zyskind, Judith W.
: APPLICANT: Wall, Daniel
: APPLICANT: Trawick, John D.
: APPLICANT: Carr, Grant J.
: APPLICANT: Yamamoto, Robert T.
: APPLICANT: Xu, H. Howard
: TITLE OF INVENTION: Identification of Essential Genes In
: TITLE OF INVENTION: Prokaryotes
: FILE REFERENCE: ELITRA 011A
: CURRENT APPLICATION NUMBER: US/09/815,242
: CURRENT FILING DATE: 2001-03-21
: PRIOR APPLICATION NUMBER: 60/191,078
: PRIOR FILING DATE: 2000-03-21
: PRIOR APPLICATION NUMBER: 60/206,848
: PRIOR FILING DATE: 2000-05-23
: PRIOR APPLICATION NUMBER: 60/207,727
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: 60/242,578
: PRIOR FILING DATE: 2000-10-23
: PRIOR APPLICATION NUMBER: 60/253,625
: PRIOR FILING DATE: 2000-11-27
: PRIOR APPLICATION NUMBER: 60/257,931
: PRIOR FILING DATE: 2000-12-22
: PRIOR APPLICATION NUMBER: 60/269,308
: PRIOR FILING DATE: 2001-02-16
: NUMBER OF SEQ ID NOS: 14110
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 13233
: LENGTH: 693
: TYPE: PRT
: ORGANISM: Streptococcus pneumoniae
US-09-815-242-13233

Query Match          9.3%: Score 68; DB 10; Length 693;
Best Local Similarity 21.7%: Pred. NO. 74;
Matches 23: Conservative 14; Mismatches 29; Indels 40; Gaps 4

OY 26 GDGSGMCLSSGOLDMLTAHGTVNVSSFSMGNALNPDNYLISKDVATGATRKVYYPVN 85
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 616 GDVGMGHVYARGRGDVGMEAHGNSQI-----VRAVYVFLA 648
    : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

```

```
OY      86 DCFPGDHVAMISKTGTNAGDTTVF---EETPNGI-----NKG 121
Db      649 EMF---GATYVLRASASQGRGTFMVFEDYHEDVPKSVQEEIIRKNKG 691

RESULT 8
US-10-125-692-18
; Sequence 18, Application US/10125692
; Publication No. US2003004429A1
; GENERAL INFORMATION:
; APPLICANT: Aderem, Alan
; APPLICANT: Hayashi, Fumitaka
; APPLICANT: Smith, Kelly D.
; APPLICANT: Underhill, David M.
; APPLICANT: Ozolsky, Adrian
; TITLE OF INVENTION: Toll-Like Receptor 5 Ligands and Methods
; FILE REFERENCE: P-TS 5155
; CURRENT APPLICATION NUMBER: US/10/125,692
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: US 60/285,477
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 506
; TYPE: PRT
; ORGANISM: S. typhimurium2
US-10-125-692-18

Query Match          9.2%: Score 67.5; DB 9; Length 506;
Best Local Similarity 27.3%: Pred. No. 57;
Matches 36; Conservative 14; Mismatches 53; Indels 29; Gaps 6;

OY      20 TTIDADGDEGMLCLSSGQLDMLTAHGSTNVVSSFSMGMALNPN---YLISKDVTGAT 76
Db      196 TTLDVSGLDLDAI-----KATGCTNGTASVTGAVFEDADNNKTFYIIGFTGAD 246

OY      77 KVVY-YYPVNDGFGD-HYAVMISKTGTNAGDT-----VVEETPNGINKKG 122
Db      247 AANGDDEVAVVADGTVTLAAGATKTPMAGATTKTEVQELKDPVAVVSADAKNALLGG 306

OY      123 ARFGLSTEANGA 134
Db      307 VD---ATDANGA 315

RESULT 9
US-09-880-748-1745
; Sequence 1745, Application US/09880748
; Publication No. US2003005937A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PF523
; CURRENT APPLICATION NUMBER: US/09/880,748
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 3239
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1745
; LENGTH: 256
; TYPE: PRT
```

```
; ORGANISM: Homo sapiens
US-09-880-748-1745

Query Match          9.2%: Score 67; DB 9; Length 256;
Best Local Similarity 19.7%: Pred. No. 28;
Matches 31; Conservative 21; Mismatches 49; Indels 56; Gaps 7;

OY      16 PAEWTTIDADGDEGMLCLSSGQL--DWLTAHGSTNVVSSFSWNG----- 58
Db      14 PSEFLSLTCTYSGS---ISSGNTYMSVNRQHPG---KGLEWIGIYIDIGNTYNSLKS 66

OY      59 ---MALNPDNYLISKDVTGAT-----KVVYYPVNDGFGDHY-----AV 95
Db      67 RVWMSVDTSKNQSFLELTSVTAADTAAYVYCARVPYDTSGCYLGEYYGMDWGQGTLY 126

OY      96 MISKTGTNAGDFTVVEETPNGINKGARFGLSTEAN 132
Db      127 TVSSGGGSGG-----GSGSGGGSQSVLTQPAS 154

RESULT 10
US-09-797-862-33
; Sequence 33, Application US/09797862
; Patent No. US20020102276A1
; GENERAL INFORMATION:
; APPLICANT: PEAK, IAN RICHARD ANSELM
; APPLICANT: JENNINGS, MICHAEL PAUL
; APPLICANT: MOXON, E. RICHARD
; TITLE OF INVENTION: NOVEL SURFACE ANTIGEN
; FILE REFERENCE: 065064/0134
; CURRENT APPLICATION NUMBER: US/09/797,862
; CURRENT FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: PCT/AU98/01031
; PRIOR FILING DATE: 1998-12-14
; PRIOR APPLICATION NUMBER: GB 9726398.2
; PRIOR FILING DATE: 1997-12-12
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 33
; LENGTH: 2353
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-09-797-862-33

Query Match          9.2%: Score 67; DB 10; Length 2353;
Best Local Similarity 25.4%: Pred. No. 43e+02;
Matches 35; Conservative 12; Mismatches 55; Indels 36; Gaps 7;

OY      20 TTIDADGDE-----GWLCLSSG---QLDWLTAHGSTNVVSSFSWNG 58
Db      1603 TVETDQKDEGNGLVAKAVIDAVNRKAGVRKTTGANGQNDDFATVASGVNTVFA-DGNG 1661

OY      59 MALNPDNYLISKDVTGATKVVYYPVNDGFGHYAVMISKTGTNAGDFTVVEETPNGI 118
Db      1662 TTAE-----VTKANDSITVKKVNVKVAADGLKDGKIDKADTTVTVADGKVT--ADN-- 1711

OY      119 NKGARF-----GLSTEAN 132
Db      1712 NGDGKKFVDASGLDALN 1729

RESULT 11
US-10-112-527-3
; Sequence 3, Application US/10112527
; Patent No. US20020168722A1
; GENERAL INFORMATION:
; APPLICANT: Gfieninger, Gerd
; APPLICANT: Applegate, Dianne
; APPLICANT: Stoike-Steben, Iara
; TITLE OF INVENTION: No. US20020168722A1e1 Cleaved Fragments of Fibrinogen
; FILE REFERENCE: Docket 454-24 CON
; CURRENT APPLICATION NUMBER: US/10/112,527
; CURRENT FILING DATE: 2002-03-29
```

;; PRIOR APPLICATION NUMBER: US 09/373,157
;; PRIOR FILING DATE: 1999-08-12
;; NUMBER OF SEQ ID NOS: 7
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 3
;; LENGTH: 221
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-112-527-3

Query Match 9.0%; Score 66; DB 9; Length 221;
Best Local Similarity 22.9%; Pred. No. 30;
Matches 24; Conservative 16; Mismatches 41; Indels 24; Gaps 4;

OY 2 DFTFFESTHGEAPAEWTTIDADGCGWLCSSGQDLMTAHGNTNVSSFSNMGMAL 61
DB 39 NFNRWODYKRG-----FGSLNDEGEFEW--LGNDYHLHTLORGSVLRLEEDMAG--- 88

OY 62 NPDNYLSKDVGTGATKYYKYPVNDGFPDGHYAVMISKTGTNAGD 106
DB 89 -----NEAVAEYHFRV--GSEAEGYALQVSYEGTAGD 119

RESULT 12
US-10-112-527-2
;; Sequence 2, Application US/10112527
;; Patent No. US20020168722A1
;; GENERAL INFORMATION:
;; APPLICANT: Griening, Gerd
;; APPLICANT: Applegate, Dianne
;; APPLICANT: Stolke-Steben, Lara
;; TITLE OF INVENTION: No. US20020168722A1 Cleaved Fragments of Fibrinogen
;; FILE REFERENCE: Docket 454-24 CON
;; CURRENT APPLICATION NUMBER: US/10/112,527
;; PRIOR FILING DATE: 2002-03-29
;; PRIOR APPLICATION NUMBER: US 09/373,157
;; PRIOR FILING DATE: 1999-08-12
;; NUMBER OF SEQ ID NOS: 7
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 2
;; LENGTH: 236
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-112-527-2

Query Match 9.0%; Score 66; DB 9; Length 236;
Best Local Similarity 22.9%; Pred. No. 32;
Matches 24; Conservative 16; Mismatches 41; Indels 24; Gaps 4;

OY 2 DFTFFESTHGEAPAEWTTIDADGCGWLCSSGQDLMTAHGNTNVSSFSNMGMAL 61
DB 54 NFNRWODYKRG-----FGSLNDEGEFEW--LGNDYHLHTLORGSVLRLEEDMAG--- 103

OY 62 NPDNYLSKDVGTGATKYYKYPVNDGFPDGHYAVMISKTGTNAGD 106
DB 104 -----NEAVAEYHFRV--GSEAEGYALQVSYEGTAGD 134

RESULT 13
US-10-112-527-1
;; Sequence 1, Application US/10112527
;; Patent No. US20020168722A1
;; GENERAL INFORMATION:
;; APPLICANT: Griening, Gerd
;; APPLICANT: Applegate, Dianne
;; APPLICANT: Stolke-Steben, Lara
;; TITLE OF INVENTION: No. US20020168722A1 Cleaved Fragments of Fibrinogen
;; FILE REFERENCE: Docket 454-24 CON
;; CURRENT APPLICATION NUMBER: US/10/112,527
;; PRIOR FILING DATE: 2002-03-29
;; PRIOR APPLICATION NUMBER: US 09/373,157
;; PRIOR FILING DATE: 1999-08-12
;; NUMBER OF SEQ ID NOS: 7

;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 1
;; LENGTH: 251
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-112-527-1

Query Match 9.0%; Score 66; DB 9; Length 251;
Best Local Similarity 22.9%; Pred. No. 35;
Matches 24; Conservative 16; Mismatches 41; Indels 24; Gaps 4;

OY 2 DFTFFESTHGEAPAEWTTIDADGCGWLCSSGQDLMTAHGNTNVSSFSNMGMAL 61
DB 69 NFNRWODYKRG-----FGSLNDEGEFEW--LGNDYHLHTLORGSVLRLEEDMAG--- 118

OY 62 NPDNYLSKDVGTGATKYYKYPVNDGFPDGHYAVMISKTGTNAGD 106
DB 119 -----NEAVAEYHFRV--GSEAEGYALQVSYEGTAGD 149

RESULT 14
US-10-141-531-26
;; Sequence 26, Application US/10141531
;; Publication No. US20030100743A1
;; GENERAL INFORMATION:
;; APPLICANT: Briggs, Steven P.
;; APPLICANT: del Val, Greg
;; APPLICANT: Desjarlais, John R.
;; APPLICANT: Helfetz, Peter
;; APPLICANT: Luginbuhl, Peter
;; APPLICANT: Muchhal, Umesh
;; TITLE OF INVENTION: Nucleic Acids and Proteins with Thioresdoxin Reductase Activity
;; FILE REFERENCE: A-71457-2/RET/RMS/RMK
;; CURRENT APPLICATION NUMBER: US/10/141,531
;; PRIOR FILING DATE: 2002-05-06
;; PRIOR APPLICATION NUMBER: US 60/370,609
;; PRIOR FILING DATE: 2002-04-05
;; PRIOR APPLICATION NUMBER: US 60/289,029
;; PRIOR FILING DATE: 2001-05-04
;; NUMBER OF SEQ ID NOS: 208
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 26
;; LENGTH: 283
;; TYPE: PRT
;; ORGANISM: Neurospora crassa
US-10-141-531-26

Query Match 9.0%; Score 66; DB 9; Length 283;
Best Local Similarity 22.0%; Pred. No. 40;
Matches 27; Conservative 24; Mismatches 52; Indels 20; Gaps 5;

OY 12 HGEAPAEWTT--IDADGCGWLC-----LSSGQDLMTAHGNTNVSSFSNMGMAL 58
DB 12 HEKVTVAFNVEVAGVGGDDGSLMSHLVYKQVTTGKRETELEANLFFAIGHDPATALYKGO 71

OY 59 MALNPDNYLSKDVGTGATKYYKYPVNDGFPDGHYAVMISKTGTNAGD 112
DB 72 LETDADGCVVTKRGTTLTVEGVFAAGD-VQDKRYROAITSAETGCMALDAEKFLSEHE 130

OY 113 ETP 115
DB 131 ETP 133

RESULT 15
US-10-141-531-68
;; Sequence 68, Application US/10141531
;; Publication No. US20030100743A1
;; GENERAL INFORMATION:
;; APPLICANT: Briggs, Steven P.
;; APPLICANT: del Val, Greg

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OK protein - protein search, using sw model

Run on: June 26, 2003, 00:46:24 ; Search time 22.0706 Seconds
(without alignments) 583.673 Million cell updates/sec

Title: US-09-980-370-6

Sequence: 1 ADFTFESTSTHGEAPAEWTTIDADGSGQGLSSQGLDMVLAHGTNNVSSFSNCGMA 134

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1: *
2: p1r2: *
3: p1r3: *
4: p1r4: *

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	716	98.1	1704	2 A55426	gingipain R (EC 3.4.22.37) precursor - Porphyromonas gingivalis
2	715	97.9	1732	2 T30836	lysine-specific cysteine proteinase (EC 3.4.22.1) - Porphyromonas gingivalis
3	711	97.4	2628	2 T28651	hemagglutinin A - Porphyromonas gingivalis
4	706	96.7	1526	2 S49763	gingipain R (EC 3.4.22.37) precursor - Porphyromonas gingivalis
5	90.5	12.4	218	2 B81094	isomerase, probable
6	90.5	12.4	286	2 D75217	hypothetical protein
7	89	12.2	418	2 T00154	hypothetical protein
8	88.5	12.1	218	2 C81847	probable hydrolase
9	83.5	11.4	352	2 A46564	thermolysin (EC 3.4.22.1)
10	83	11.4	372	2 B70819	probable carnitine
11	82.5	11.3	336	2 A99283	hypothetical protein
12	81.5	11.2	386	2 G69189	hypothetical protein
13	80	11.0	428	2 H63767	cytoplasmic protein
14	79.5	10.9	548	1 HYRSP	cellulase (EC 3.2.1.4)
15	78.5	10.8	551	2 B3706	heparin lyase (EC 3.2.1.16)
16	78.5	10.8	346	1 B4509	aldehyde lyase
17	78.5	10.8	453	2 A46022	tryptophanase (EC 3.2.1.17)
18	78.5	10.8	499	2 A27198	cellulase (EC 3.2.1.4)
19	78	10.7	527	2 PNO114	microbial metalloprotease
20	76	10.4	187	2 F64834	limbrial-like protein
21	75.5	10.3	439	2 T02956	alpha-amylase (EC 3.2.1.1)
22	75.5	10.3	607	2 S49528	endoxyalanase - rum
23	75.5	10.3	1034	2 A36108	integrin alpha-V
24	75.5	10.3	759	2 A49398	cycloartenol synthase
25	75	10.3	985	2 T29910	hypothetical protein
26	74.5	10.2	479	2 A69832	probable acid-CoA
27	74.5	10.2	548	1 HYRSP	heparin lyase (EC 3.2.1.16)
28	74.5	10.2	550	1 T36746	probable serine/threonine protease

30	74	10.1	453	2 B49022	tryptophanase (EC 3.2.1.17)
31	74	10.1	504	2 S54744	cellulase (EC 3.2.1.4)
32	74	10.1	354	2 G74258	hypothetical protein
33	74	10.1	239	2 T28651	hemagglutinin A - Porphyromonas gingivalis
34	73.5	10.1	544	2 A42464	microbial metalloprotease
35	73.5	10.1	546	2 J04113	neutral protease
36	73.5	10.1	546	2 S72176	thermolysin (EC 3.4.22.1)
37	73.5	10.1	546	2 S72176	thermolysin (EC 3.4.22.1)
38	73.5	10.1	546	2 A14423	beta-glucosidase
39	73.5	10.1	4936	2 A14255	hypothetical protein
40	73.5	10.1	435	2 S12625	alpha-amylase (EC 3.2.1.1)
41	73	10.0	435	2 J07137	isochlorismate synthase
42	73	10.0	441	2 E84264	flagellar hook protein
43	72.5	9.9	264	2 G69622	extracellular protease
44	72.5	9.9	449	2 J07306	extracellular protease
45	72.5	9.9	449	2 J07306	extracellular protease

ALIGNMENTS

RESULT 1
A55426
gingipain R (EC 3.4.22.37) precursor - Porphyromonas gingivalis
N:Alternate names: 50K high molecular mass arginine-specific cysteine proteinase; HGP
C:Species: Porphyromonas gingivalis
C>Date: 10-Feb-1995 #sequence_revision 10-Feb-1995 #text_change 08-Oct-1999
G:Accession: A55426; D53113
R:Protein: N.; Potempa, J.; Pike, R.N.; Prochazka, V.; Kiefer, M.C.; Travis, J.; Barr, J.
J: Biol. Chem. 270, 1007-1010, 1995
A:Title: Molecular cloning and structural characterization of the Arg-gingipain protease
A:Reference number: A55426; MIM:9518080; PMID:753531
A:Accession: D53113
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-1704 <PAV>
A:Cross-references: GB:U15282; NID:9557067; PID:AAA69539.1; PID:9557068
R:Pike, R.; McGraw, W.; Potempa, J.; Travis, J.
J: Biol. Chem. 269, 406-411, 1994
A:Title: Lysine- and arginine-specific proteinases from Porphyromonas gingivalis. Iso
A:Reference number: A53113; MIM:94103245; PMID:8276827
A:Accession: D53113
A:Status: preliminary
A:Molecule type: protein
A:Residues: 228-249 <PIK>
A:Experimental source: H66
A:Note: sequence extracted from NCBI database (NCBI:141694)
C:Keywords: cysteine proteinase; hydrolase
Query Match 98.1% Score 716; DB 2; Length 1704;
Best Similarity 97.8% Pred. No. 9.2e-58;
Matches 131; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
OY 1 ADFTFESTSTHGEAPAEWTTIDADGSGQGLSSQGLDMVLAHGTNNVSSFSNCGMA 60
DB 1137 ADFTFESTSTHGEAPAEWTTIDADGSGQGLSSQGLDMVLAHGTNNVSSFSNCGMA 1196
OY 61 LNDPNVLSKVGATKRYKYPYVPCQDHYANISKTCNMADDFVFEETPNCNK 120
DB 1197 LNDPNVLSKVGATKRYKYPYVPCQDHYANISKTCNMADDFVFEETPNCNK 1256
OY 121 GGAPELSTANCA 134
DB 1257 GGAPELSTANCA 1270
RESULT 2
T30836
lysine-specific cysteine proteinase porphyrin (EC 3.4.22.-) - Porphyromonas gingivalis
N:Alternate names: lysine-specific cysteine proteinase 1, 60K
C:Species: Porphyromonas gingivalis
C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 17-Nov-2000
C:Accession: T30836; T30837; T30526; A53113

R:Barbey-Gallagher, C.A.; Han, N.; Patel, J.M.; Whitlock, J.; Proguliske-Fox, A.; Lantz, J. Infect. Immun. 178:2734-2741, 1996
 A:Title: Analysis of the prpC gene encoding porphyrin, a cysteine proteinase of Porphyromonas gingivalis
 A:Reference number: 220895; MUID:96213011; PMID:663659
 A:Accession: T30836
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1732 <BAR>
 A:Cross-references: EMBL:U42210; NID:91314325; PID:91314326; PIDN:AA06565.1
 R:Slakoski, N.; Cleal, S.M.; Reynolds, E.C.
 A:Submitted to the EMBL Data Library, October 1996
 A:Reference number: 220896
 A:Accession: T30837
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-795; 1,797-1289; 1,1391-1478; 1,1480-1732 <SLA>
 A:Cross-references: EMBL:075366; NID:92182811; PID:92182812; PIDN:AA060809.1
 R:Leal, D.F.; Bact 1035-1042, 1998
 A:Title: IS195, an insertion sequence element associated with protease genes in Porphyromonas gingivalis
 A:Reference number: 220844; MUID:98298016; PMID:9632563
 A:Accession: T30526
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1350; 1,352-1363; 1,365-1447; 1,1449-1732 <LEM>
 A:Cross-references: EMBL:AF017059; NID:92738802; PID:92738803; PIDN:AA026523.1
 R:Pike, R.; McGraw, W.; Potempa, J.; Travis, J.
 A:Biochem. J. 269: 406-411, 1994
 A:Title: Lysine- and arginine-specific proteinases from Porphyromonas gingivalis. Isolation and characterization
 A:Reference number: A53113; MUID:94103245; PMID:8276827
 A:Accession: A53113
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 229-249 <PIK>
 A:Experimental source: H66
 A:Keywords: cysteine proteinase; hydrolyase
 C:Keywords: cysteine proteinase; hydrolyase
 Query Match 97.9%; Score 715; DB 2; Length 1732;
 Best Local Similarity 97.8%; Pred. No. 1.2e-57;
 Matches 131; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 Oy 1 ADETFEFSSSTHGEAPAEWTTIDADCGGGMCLSSGOLDMLTAHAGTNNVVSFSWNGMA 60
 Db 1157 ADETFEFSSSTHGEAPAEWTTIDADCGGGMCLSSGOLDMLTAHAGTNNVVSFSWNGMA 1216
 Oy 61 LNDPNTLISKDVGTGATKVKYVYVNDGFPDGHVAVMISTGTNAAGDTVVFEETPNCINK 120
 Db 1217 LNDPNTLISKDVGTGATKVKYVYVNDGFPDGHVAVMISTGTNAAGDTVVFEETPNCINK 1276
 Oy 121 GGAARFGLSTENGA 134
 Db 1277 GGAARFGLSTENGA 1290

RESULT 3
 T28651
 hemagglutinin A - Porphyromonas gingivalis
 C:Species: Porphyromonas gingivalis
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 03-Aug-2001
 C:Accession: T28651
 R:Han, N.; Whitlock, J.; Proguliske-Fox, A.
 A:Infect. Immun. 64: 4000-4007, 1996
 A:Title: The hemagglutinin gene A (haga) of Porphyromonas gingivalis 381 contains four
 A:Reference number: 220494; MUID:97047672; PMID:8926061
 A:Accession: T28651
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-2628 <HAN>
 A:Cross-references: EMBL:U41807; NID:91552410; PID:91469916; PIDN:AA017128.1
 C:Keywords:

A:Gene: haga
 Query Match 97.4%; Score 711; DB 2; Length 2628;
 Best Local Similarity 97.0%; Pred. No. 4.4e-57;
 Matches 130; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 Oy 1 ADETFEFSSSTHGEAPAEWTTIDADCGGGMCLSSGOLDMLTAHAGTNNVVSFSWNGMA 60
 Db 685 ADETFEFSSSTHGEAPAEWTTIDADCGGGMCLSSGOLDMLTAHAGTNNVVSFSWNGMA 744
 Oy 61 LNDPNTLISKDVGTGATKVKYVYVNDGFPDGHVAVMISTGTNAAGDTVVFEETPNCINK 120
 Db 745 LNDPNTLISKDVGTGATKVKYVYVNDGFPDGHVAVMISTGTNAAGDTVVFEETPNCINK 804
 Oy 121 GGAARFGLSTENGA 134
 Db 805 GGAARFGLSTENGA 818

RESULT 4
 S49763
 gingipain B (PGC 3.4.22.37) precursor - Porphyromonas gingivalis (fragment)
 C:Species: Porphyromonas gingivalis
 C>Date: 05-Mar-1995 #sequence_revision 12-May-1995 #text_change 31-Mar-1997
 C:Accession: S49763
 R:Aduse-Opoku, J.; Muir, J.; Slaney, J.M.; Rangarajan, M.; Curtis, M.A.
 A:Submitted to the EMBL Data Library, November 1994
 A:Description: Cloning, sequence analysis and expression in Escherichia coli of prpR1
 A:Reference number: S49763
 A:Accession: S49763
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1526 <ADU>
 A:Cross-references: EMBL:X62860
 C:Keywords: prpR1
 C:Keywords: cysteine proteinase; hydrolyase
 Query Match 96.7%; Score 706; DB 2; Length 1526;
 Best Local Similarity 97.0%; Pred. No. 6.8e-57;
 Matches 130; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 Oy 1 ADETFEFSSSTHGEAPAEWTTIDADCGGGMCLSSGOLDMLTAHAGTNNVVSFSWNGMA 60
 Db 1139 ADETFEFSSSTHGEAPAEWTTIDADCGGGMCLSSGOLDMLTAHAGTNNVVSFSWNGMA 1198
 Oy 61 LNDPNTLISKDVGTGATKVKYVYVNDGFPDGHVAVMISTGTNAAGDTVVFEETPNCINK 120
 Db 1199 LNDPNTLISKDVGTGATKVKYVYVNDGFPDGHVAVMISTGTNAAGDTVVFEETPNCINK 1258
 Oy 121 GGAARFGLSTENGA 134
 Db 1259 GGAARFGLSTENGA 1272

RESULT 5
 B81094
 isomerase, probable NMB138 (imported) - Neisseria meningitidis (strain MC58 serogroup
 C:Species: Neisseria meningitidis
 C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
 C:Accession: B81094
 R:Rettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, R.; Hickey, E.R.; Haft, D.H.; Salberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.; H.; Qiu, H.; Yamashiro, J.; Gill, J.; Scarlato, V.; Messigman, V.; Pizzi, M.
 A:Science 287: 1809-1815, 2000
 A:Authors: Grundl, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappaport, R.
 A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
 A:Reference number: A81000; MUID:20175755; PMID:10710307
 A:Accession: B81094
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-218 <RET>
 A:Cross-references: GB:AE002482; GB:AE002098; NID:97226577; PIDN:AA01713.1; PID:9722

A:Experimental source: serogroup B, strain MC58

C:Genetics:

A:Gene: NMB1338

C:Superfamily: 2-hydroxyhepta-2,4-diene-1,7-dioate isomerase

Query Match 12.4%; Score 90.5; DB 2; Length 218;

Best Local Similarity 26.6%; Pred. No. 0.42;

Matches 38; Conservative 14; Mismatches 48; Indels 43; Gaps 6;

23 DADGDEG-----WLCSSGQDMLTAHG--GNNVSSFSNMG 59

Db 77 DSDGTGEGKDIICVAGYGVGLDTARDIQCRKEKGLPWLAKGFRHSACVSPFAAGR 136

Qy 60 ALNPDNYLISKDVYGTATKVK-----YYPVNDGFPEDHYAVMISKT---GTNAGDFTV 110

Db 137 IGNEPEKLVFLKONGVYKQGGDTGLMIYPIRE-----ILHKLADYGLGKGD--LV 185

Qy 111 FEETPNKNGKARFGLSTEANG 133

Db 186 FTGTSGVGAIGAGDMLALELDG 208

RESULT 6

D75217

probable 2-acetyl-1-alkylglycerophosph ocholone esterase PAB2176 - Pyrococcus abyssi (st

C:Species: Pyrococcus abyssi

C>Date: 20-Aug-1999 #sequence_rev1sion 20-Aug-1999 #text_change 20-Aug-1999

C:Accession: D75217

R:Anonymous, Genoscope

submitted to the EMBL Data Library, July 1999

A:Description: Pyrococcus abyssi genome sequence: Insights into archaeal chromosome stru

A:Reference number: A75001

A:Accession: D75217

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-286 <RAW>

A:Cross-references: GB:AJ248283; GB:AL096836; NID:g5457433; PIDN:CAB49187.1; PID:el51508

A:Experimental source: strain Orsay

C:Genetics:

A:Gene: PAB2176

Query Match 12.4%; Score 90.5; DB 2; Length 286;

Best Local Similarity 27.9%; Pred. No. 0.58; Mismatches 26; Indels 33; Gaps 5;

Matches 29; Conservative 16; Mismatches 26; Indels 33; Gaps 5;

Qy 6 TFESSTHGAPAEVTTIDADGDEGMLCSSLGQDMLTAHGNTNVSSFSMN----- 57

Db 102 TDFPRAHGESEGSKTIT---GDKE--ILDLSGAIIDWLLSNTYTKITALIGFSMGAMVTR 156

Qy 58 -----GMLNPDNYLISKDVYGTATKVKY-----YYPV 84

Db 157 ALAEDERVCCGIADSPPIYI---DKTGAGKTKIFANLPFLPI 197

RESULT 7

T00154

hypothetical protein 21 - Staphylococcus aureus phage phi PVL

C:Species: Staphylococcus aureus phage phi PVL

C>Date: 23-Apr-1999 #sequence_rev1sion 23-Apr-1999 #text_change 21-Jul-2000

C:Accession: T00154

R:Kaneko, Y.; Kimura, T.; Kawakami, Y.; Tomita, T.; Kamio, Y.

Bioest. Biotechnol. Biochem. 61, 1960-1962, 1997

A:Title: Pantone-Valentine leukocidin genes in a phage-like particle isolated from mitomy

A:Reference number: Z14119; MUID:98067870; PMID:9404084

A:Accession: T00154

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-418 <RAW>

A:Cross-references: EMBL:AB009666; NID:g3341907; PIDN:BA431894.1; PID:g3341928

C:Superfamily: Staphylococcus aureus phage phi PVL hypothetical protein 21

Query Match

Best Local Similarity 12.2%; Score 89; DB 2; Length 418;

Matches 36; Conservative 19; Mismatches 55; Indels 22; Gaps 7;

Qy 20 TTIDADGDEGMLCSSLGQDMLTAHGNTNVSSFSNMG--MLNPDNYLISKDVYGTGA 75

Db 94 TYIDGEGDEG-----SSGITQWMDKTYSDSGNMGITINSYGVAVLTSNRRVYLESTAS 149

Qy 76 TKVK-----YYPVNDGFPG--DHYAVMISKT--GTNAGDFTVFEETPNKNGKARF-- 125

Db 150 SNIKSKQAPVYLYLPNDKVPGLNRFATLSNADNAYSSDGYIMFGSDENYDYGAGIRFSK 209

Qy 126 ----GLSTEANG 133

Db 210 ERNKGVLQVYNG 221

RESULT 8

C81847

probable hydrolase NMA1552 [imported] - Neisseria meningitidis (strain Z2491 serogrou

C:Species: Neisseria meningitidis

C>Date: 05-May-2000 #sequence_rev1sion 05-May-2000 #text_change 02-Feb-2001

C:Accession: C81847

R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mo

; Holroyd, S.; Jørgensen, K.; Leather, S.; Mouton, S.; Mungall, K.; Quail, M.A.; Rajandre

Nature 404, 502-506, 2000

A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491

A:Reference number: A81775; MUID:20222556; PMID:10761919

A:Accession: C81847

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-218 <PAR>

A:Cross-references: GB:AL162756; GB:AL157959; NID:g7380091; PIDN:CAB84779.1; PID:g738

A:Experimental source: serogroup A, strain Z2491

C:Genetics:

A:Gene: NMA1552

C:Superfamily: 2-hydroxyhepta-2,4-diene-1,7-dioate isomerase

Query Match 12.1%; Score 88.5; DB 2; Length 218;

Best Local Similarity 26.6%; Pred. No. 0.65; Mismatches 48; Indels 43; Gaps 6;

Matches 38; Conservative 14; Mismatches 48; Indels 43; Gaps 6;

Qy 23 DADGDEG-----WLCSSGQDMLTAHG--GNNVSSFSNMG 59

Db 77 DSDGTGEGKDIICVAGYGVGLDTARDIQCRKEKGLPWLAKGFRHSACVSPFAAGR 136

Qy 60 ALNPDNYLISKDVYGTATKVK-----YYPVNDGFPEDHYAVMISKT---GTNAGDFTV 110

Db 137 IGNEPEKLVFLKONGVYKQGGDTGLMIYPIRE-----ILHKLADYGLGKGD--LV 185

Qy 111 FEETPNKNGKARFGLSTEANG 133

Db 186 FTGTSGVGAIGAGDMLALELDG 208

RESULT 9

A46564

thermolysin (EC 3.4.24.27) homolog precursor - Bacillus stearothermophilus

C:Species: Bacillus stearothermophilus

C>Date: 03-Mar-1994 #sequence_rev1sion 03-Mar-1994 #text_change 22-Jun-1999

C:Accession: A46564

R:Kubo, M.; Imanaka, T.

J. Gen. Microbiol. 134, 1883-1892, 1988

A:Title: Cloning and nucleotide sequence of the highly thermostable neutral protease

A:Reference number: A46564; MUID:89235633; PMID:3149972

A:Accession: A46564

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-552 <KUB>

A:Cross-references: GB:M21663; NID:g143746; PIDN:AA802774.1; PID:g143747

A>Note: the authors translated the codon TCG for residue 32 as Thr, CTG for residue 5

C:Superfamily: thermolysin

C:Keywords: hydrolase; metalloproteinase; zinc

F:378,382,402/Binding site: zinc (His, His, Glu) #status predicted

F:379,467/Active site: Glu, His #status predicted

C:Genetics:
A:Gene: AT29707050

A:Map position: 2

C:Superfamily: yeast lanosterol synthase

Query Match 11.0%; Score 80; DB 2; Length 759;

Best Local Similarity 28.7%; Pred. No. 17;

Matches 25; Conservative 13; Mismatches 15; Indels 34; Gaps 7;

QY 13 GAPAPMTTIDADGEGWLCSSGQDMLTAHAGC-TNVVS-----SFSWNG-MA 60

DB 188 GCGP-----NDGGD-----MEKGR-DWLNHGATNTSMGKMWLSVLGAFEMSGNNP 235

QY 61 LNPDMVLISKDYTGATKKYKYYPPVNDG 87

DB 236 LPPEITWL-----PYFLPIHPG 252

RESULT 14

503767

cellulase (EC 3.2.1.4) Z precursor - Erwina chrysanthemi

N:Alternate names: endo-1,4-beta-glucanase; extracellular endoglucanase Z

C:Species: Erwina chrysanthemi

A:Variety: strain 3937

C:Date: 21-Nov-1993 #sequence_revision 01-Dec-1995 #text_change 08-Oct-1999

C:Accession: S03767

R:Clustep: A.; Caml. B.; Aymeric, J.L.; Ball, G.; Creuzet, N.

Mol. Microbiol. 2, 159-164, 1988

A:Title: Homology between endoglucanase Z of Erwina chrysanthemi and endoglucanases of

A:Reference number: S03767; MUID:88216177; PMID:2835589

A:Accession: S03767

A:Molecule type: DNA

A:Residues: 1-428 <GUI>

A:Cross-references: EMBL:Y00540; NID:941091; PIND:CAA68604.1; PID:941092

C:Genetics:

A:Gene: celZ

C:Function: hydrolysis of 1,4-beta-D-glucosidic linkages in beta-D-glucans such as ce

A:Pathway: cellulose degradation

C:Keywords: glycosidase; hydrolase; polysaccharide degradation

F:1-43/Domain: signal sequence #status predicted <SIG>

F:44-428/Product: cellulase Z #status predicted <MAT>

Query Match 10.9%; Score 79.5; DB 2; Length 428;

Best Local Similarity 28.7%; Pred. No. 9.5; Indels 31; Gaps 6;

Matches 29; Conservative 15; Mismatches 26; Indels 31; Gaps 6;

QY 3 FTEFESTHGEA-----PAEWTIDADGEGWLCSSGQD-MLTAH 45

DB 232 YLHMFYAGHGESLNNKARQALNGLALFVEMGTNVNDGNG-----VNOTETDAWTFM 287

QY 46 GGTNVSSFSNMGMLNDPN-----YLISKDYT-GATKKV 79

DB 288 RDNNTIOLTONM---ALNDKNEGASTYIPDSKNLIESGKKV 325

RESULT 15

HYBST

thermolysin (EC 3.4.24.27) precursor [validated] - Bacillus "thermoproteolyticus"

N:Alternate names: Bacillus thermoproteolyticus neutral proteinase; microbial metallopro

C:Species: Bacillus "thermoproteolyticus"

C:Date: 24-Apr-1984 #sequence_revision 07-Jun-1996 #text_change 15-Sep-2000

C:Accession: I40579; A00993; S41312

R:O'Donohue, M.J.; Roques, B.P.; Beaumont, A.

Biochem. J. 300, 599-603, 1994

A:Title: Cloning and expression in Bacillus subtilis of the npr gene from Bacillus therm

A:Reference number: I40579; MUID:94271180; PMID:8002967

A:Accession: I40579

A:Molecule type: DNA

A:Residues: 1-408, 'K', 410-548 <OAD>

A:Cross-references: EMBL:X76865; NID:9441266; PIND:CAA54291.1; PID:9441267

A:Note: submitted to the EMBL Data Library, December 1993

R:Itani, K.; Hermodson, M.A.; Ericsson, L.H.; Walsh, K.A.; Neurath, H.

Nature New Biol. 238, 35-37, 1972

A:Title: Amino-acid sequence of thermolysin.

A:Reference number: A00993

A:Accession: A00993

A:Molecule type: Protein

A:Residues: 233-268, 'D', 270-350, 'E', 352-399, 'I', 401-548 <TIT>

R:Matthews, B.W.; Holmes, M.A.

submitted to the Brookhaven Protein Data Bank, February 1982

A:Reference number: A50631; PDB:3TLN

A:Contents: annotation; X-ray crystallography, 1.6 angstroms, residues 233-268, 'D', 27

J. Mol. Biol. 160, 623-639, 1982

A:Title: Structure of thermolysin redefined at 1.6 A resolution.

A:Reference number: A92888; MUID:83085812; PMID:7175940

A:Contents: annotation; X-ray crystallography, 1.6 angstroms

R:Matthews, B.W.; Weaver, L.H.; Kester, W.R.

J. Biol. Chem. 249, 8030-8044, 1974

A:Title: The conformation of thermolysin.

A:Reference number: A92165; MUID:75041142; PMID:4214815

A:Contents: annotation; X-ray crystallography, 2.3 angstroms

R:Burstein, Y.; Walsh, K.A.; Neurath, H.

Biochemistry 13, 205-210, 1974

A:Title: Evidence of an essential histidine residue in thermolysin.

A:Reference number: A90377; MUID:74052951; PMID:4808703

A:Contents: annotation; active site

C:Function:

A:Description: metalloendopeptidase preferentially hydrolyzes peptide bonds on the am

C:Superfamily: thermolysin

C:Keywords: calcium; extracellular protein; hydrolase; metalloproteinase; zinc

F:1-18/Domain: signal sequence #status predicted <SIG>

F:19-232/Domain: propeptide #status predicted <PRO>

F:233-548/Product: thermolysin #status experimental <MAT>

F:289,291,293/Binding site: calcium (Asp, Asp, Glu) #status experimental

F:379,409,417,419,422/Binding site: calcium (Asp, Glu, Glu, Asp, Glu, Glu) #status experim

F:374,378,398/Binding site: zinc (His, His, Glu) #status experimental

F:375,463/Active site: Glu, His #status predicted

F:409,415,417,422/Binding site: calcium (Glu, Asn, Asp, Glu) #status experimental

F:425,426,429,432/Binding site: calcium (Tyr, Thr, Ile, Asp) #status experimental

Query Match 10.9%; Score 79.5; DB 1; Length 548;

Best Local Similarity 24.7%; Pred. No. 13;

Matches 40; Conservative 28; Mismatches 53; Indels 41; Gaps 10;

QY 9 SSTRH---GEAPAWT---TIDADGEGWLCSSGQDMLTAHAGTNNVSSFSNMGMLN 62

DB 334 SSVHYSGVYNNAPMNGSCOWVYGDGQTFPIPLSG-ID-VVAHELTHAVTDYT-AGLIYQ 390

QY 63 PDNVLISK---DVTGATKKYKYYPPVNDG-----PG-----DHY 93

DB 391 NESGALINEMSDIFG-TLVEFYNNKRPWEIGEDVYTPDISDSLSKMSDPKAYGDPHY 449

QY 94 AVMISKGTNAGDF---TVVFETPNCINKGARFGLSTEANG 133

DB 450 SKRYTGTDNGGVHINSGLINKAAVYLISOGTHYGVSVIG 491

Search completed: June 26, 2003, 00:52:22
Job time : 23.0706 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 26, 2003, 00:44:19 ; Search time 14.1882 Seconds
(without alignments)
391.721 Million cell updates/sec

Title: us-09-980-370-6

Perfect score: 730
Sequence: 1 ADPTEPESSTHGEAPAEWT.....PNGINKGARGLSTENGCA 134

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SWISSPROT_40*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	714	97.8	989	PRTH_PORGI	P46071 porphyromon
2	711	97.4	2628	HAGA_PORGI	Q51845 porphyromon
3	80	11.0	759	CASI_ARATH	P38605 arabidopsis
4	79.5	10.9	316	THRE_BACTH	P00800 bacillus th
5	79.5	10.9	551	NPBS_BACST	P43133 bacillus th
6	78.5	10.8	348	GALM_STRTT	P21935 streptococ
7	78.5	10.8	453	TNAL_SYMTT	P31014 symbiodact
8	78	10.7	426	GNUN_ERMCH	P07103 erwinia chr
9	76	10.4	527	NPBE_BACBR	P43263 bacillus br
10	75.5	10.3	171	YCBV_ECOLI	P75860 escherichia
11	75.5	10.3	1034	ITAV_CHICK	P26008 gallus gall
12	74.5	10.2	548	THRE_BACST	P06874 bacillus st
13	74	10.1	452	TNA2_SYMTT	P31015 symbiodact
14	74	10.1	504	GNUN_ERMCA	Q53395 erwinia car
15	73.5	10.1	297	THW2_MYCTU	P56888 mycobacteri
16	73.5	10.1	544	NPBE_BACCL	P23384 bacillus ca
17	73	10.0	435	AMSD_ORISA	P27933 oryza sativ
18	72.5	9.9	364	FLGG_BACSU	P23446 bacillus su
19	72	9.9	307	BLAC_MYCTU	Q10670 mycobacteri
20	72	9.9	519	VP6_MTV	P12325 wound tumor
21	72	9.9	1953	BIGA_SALTY	P25927 salmonella
22	71.5	9.8	474	SHUV_ECOLI	P09745 escherichia
23	71	9.7	505	GNUN_ERMCA	Q47095 erwinia car
24	71	9.7	519	VP6_MTV	P31610 wound tumor
25	70.5	9.7	607	XYNA_NDOPA	P29127 neocallimas
26	70	9.6	221	XYN2_EMENT	P55333 emericella
27	70	9.6	1045	GNUN_CELFT	P52225 celidomona
28	70	9.6	1045	ITAV_HUMAN	P08648 homo sapien
29	69.5	9.5	312	PGTE_SALTY	P06185 salmonella
30	69.5	9.5	547	FUMA_ECOLI	P00923 escherichia
31	69.5	9.5	579	ITAV_SALTY	P40720 salmonella
32	69.5	9.5	1044	ITAV_MOUSE	P43406 mus musculu
33	69.5	9.5	1176	CIAG_BACTU	O95515 bacillus th

34	69	9.5	249	1	GPPI_YEAST	P41277 saccharomyc
35	69	9.5	463	1	STHA_PSEAE	P57112 pseudomonas
36	69	9.5	705	1	CMBA_BACSU	002113 bacillus su
37	69	9.5	958	1	TRH_DROME	Q24119 drosophila
38	68.5	9.4	321	1	CTSI_HOMAM	P25784 homarus ame
39	68.5	9.4	521	1	NPBE_BACSU	P06142 bacillus su
40	68.5	9.4	1048	1	ITAV_HUMAN	P06756 homo sapien
41	68.5	9.4	3672	1	LMR2_CAEEL	Q21313 caenorhabdi
42	68	9.3	250	1	GPPI_YEAST	P40106 saccharomyc
43	68	9.3	373	1	RF2_AOUAE	Q67695 aquifex aeo
44	68	9.3	444	1	GNUN_ERMCA	O59394 erwinia car
45	68	9.3	523	1	TYD5_PAPSO	P54771 papaver som

ALIGNMENTS

RESULT 1
ID PRTH_PORGI STANDARD: PRT: 989 AA.
AC P46071:
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Protease prth (EC 3.4.22.-).
GN PRTH.
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteriia; Bacteroidetes; Bacteroides; Bacteroidales;
OC Porphyromonadaceae; Porphyromonas.
OX NCBI_TaxID=837;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M83;
RX MEDLINE=95012612; PubMed=7927685;
RA Fletcher H.M., Schenkein H.A., Macrina F.L.;
RT Cloning and characterization of a new protease gene (prth) from
RT Porphyromonas gingivalis.";
RL Infect. Immun. 62:4279-4286(1994).
RN [2]
RP ERRATUM.
RA Fletcher H.M., Schenkein H.A., Macrina F.L.;
RL Infect. Immun. 62:5707-5707(1994).
CC -1- FUNCTION: CLEAVES HUMAN COMPLEMENT COMPONENT C3. MAY ENABLE
CC P. GINGIVALIS TO EVADE COMPLEMENT-MEDIATED KILLING DURING THE
CC IMMUNE RESPONSE. PLAYS AN IMPORTANT ROLE IN SOFT TISSUE INFECTIONS
CC AND IS A VIRULENCE FACTOR.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C25.
CC -1- SUBCELLULAR LOCATION: IN MEMBRANE VESICLES.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: L27483; AA051298.1; -
CC DR HSBP; P23882; IEMT.
CC DR MEROPS; C25.001; -
CC DR InterPro: IPR001769; Peptidase_C25.
CC DR InterPro: IPR002376; formyl_transf.
CC Pfam: PF00551; formyl_transf. 1.
CC DR Pfam: PF01364; Peptidase_C25; 1.
CC KW Hydrolyase; Thiol protease; Repeat; Virulence.
CC REPEAT
CC REPEAT 270 323
CC REPEAT 528 581
CC SEQUENCE 989 AA; 110238 MW; F885FE8A3AC8944C CRC64;
Query Match 97.8%; Score 714; DB 1; Length 989;
Best Local Similarity 97.8%; Pred. No. 6.4e-60;
Matches 131; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 ADFTETPSSSTHGEAPAEWTTIDADGEGMLSSGOLDMLTAHGTNNVSSFSMNMA 60
 DB 90 ADFTETPSSSTHGEAPAEWTTIDADGEGMLSSGOLDMLTAHGTNNVSSFSMNMA 149
 OY 61 LNPBNVLSKDVYTGATKTKYKYYYPVNDGFPDGHVAVNISTGTNAGDFTVFEETPENGINK 120
 DB 150 LNPBNVLSKDVYTGATKTKYKYYYPVNDGFPDGHVAVNISTGTNAGDFTVFEETPENGINK 209
 OY 121 GGARFGLSTEANGA 134
 DB 210 GGARFGLSTEADGA 223

RESULT 2

HAGA_PORGI STANDARD: PRT: 2628 AA.
 AC 051845:
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hemagglutinin A precursor.
 GN HAGA.
 OS Porphyromonas gingivalis (Bacteroides gingivalis).
 OC Bacteria; Bacteroidetes; Bacteroides; Bacteroidales;
 OC Porphyromonadaceae; Porphyromonas.
 OX NCBI_TaxID=837;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=381;
 RX MEDLINE=97047672; PubMed=89256061;
 RA Han N., Whitlock J., Prohnske-Fox A.;
 RT "The hemagglutinin gene A (haga) of Porphyromonas gingivalis 381
 RL contains four large, contiguous, direct repeats.";
 RL Infect. Immun. 64:4000-4007(1996).
 CC -1- FUNCTION: AGGLUTININATES ERYTHROCYTES.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C25.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: U01807; A017128.1;
 DR InterPro: IPR001769; Peptidase_C25.
 DR Pfam: PF01364; Peptidase_C25; 6.
 KW Hemagglutinin; Virulence; Hydrolase; Thiol protease; Signal; Repeat.
 FT SIGNAL
 FT CHAIN 1 24
 FT DOMAIN 25 2628 HEMAGGLUTININ A.
 FT DOMAIN 25 539 PEPTIDASE C25-LIKE 1.
 FT DOMAIN 540 995 PEPTIDASE C25-LIKE 2.
 FT DOMAIN 996 1451 PEPTIDASE C25-LIKE 3.
 FT DOMAIN 1452 1907 PEPTIDASE C25-LIKE 4.
 FT DOMAIN 2074 2628 PEPTIDASE C25-LIKE 5.
 SO SEQUENCE 2628 AA; 283324 MW; 61C4DE32540C99DA CRC64;

Query Match 97.48; Score 711; DB 1; Length 2628;
 Best Local Similarity 97.08; Pred. No. 3; e-59;

Matches 130; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 ADFTETPSSSTHGEAPAEWTTIDADGEGMLSSGOLDMLTAHGTNNVSSFSMNMA 60
 DB 685 ADFTETPSSSTHGEAPAEWTTIDADGEGMLSSGOLDMLTAHGTNNVSSFSMNMA 744
 OY 61 LNPBNVLSKDVYTGATKTKYKYYYPVNDGFPDGHVAVNISTGTNAGDFTVFEETPENGINK 120
 DB 745 LNPBNVLSKDVYTGATKTKYKYYYPVNDGFPDGHVAVNISTGTNAGDFTVFEETPENGINK 804
 OY 121 GGARFGLSTEANGA 134
 DB 805 GGARFGLSTEADGA 818

RESULT 3

CASI_ARATH

ID CASI_ARATH STANDARD: PRT: 759 AA.

AC P38605; P92967;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Cycloartenol synthase (EC 5.4.99.8) (2,3-epoxysqualene--cycloartenol
 DE cyclase).
 GN CASI OR AT2G07050 OR T4E14.16.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC euroside II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Landsberg erecta;
 RX MEDLINE=94089686; PubMed=7505443;
 RA Corey E.J., Matsuda S.P.T., Bartel B.;
 RT "Isolation of an Arabidopsis thaliana gene encoding cycloartenol
 RT synthase by functional expression in a yeast mutant lacking
 RT lanosterol synthase by the use of a chromotographic screen.";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:11628-11632(1993).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=20083487; PubMed=10617197;
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
 RA Bell C.R., Ketchum K.A., Lee J.J., Rosting C.M., Koo H.L.,
 RA Moffat K.S., Cronin L.A., Shen M., Pal G., Van Aken S., Umayam L.,
 RA Tallon L.J., Gill J.E., Adams M.D., Carrera A.J., Cressy T.H.,
 RA Goodman H.M., Somerville C.R., Copenhaver G.P., Preuss D.,
 RA Nierman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,
 RA Venter J.C.;
 RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
 RT thaliana.";
 RL Nature 402:761-768(1999).
 RN (3)
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RA Shiozaki K., Davis R.W., Ecker J.R., Theologis A.;
 RT "Riken Arabidopsis full length cDNA clones (RAFLs) sequenced by the
 RT SSP consortium (Salk/Stanford/PGEC).";
 RL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: (S)-2,3-epoxysqualene = cycloartenol.
 CC -1- SIMILARITY: BELONGS TO THE TERPENE CYCLASE/MUTASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: U02555; AAC04931.1;
 DR EMBL: AC005171; AAM15015.1;
 DR EMBL: AY094394; AAM19773.1;
 DR InterPro: IPR001330; Prenyltrans.
 DR InterPro: IPR002365; Terpene_synth.
 DR Pfam: PF00432; prenyltrans; 3.
 DR PROSITE: PS01074; TERPENE_SYNTHASES; 1.
 KW Isomerase; Repeat.
 FT REPEAT 147 188 PFTB 1.
 FT REPEAT 512 557 PFTB 2.
 FT REPEAT 589 629 PFTB 3.
 FT REPEAT 638 679 PFTB 4.
 FT REPEAT 700 741 PFTB 5.

FT CONFLICT 501 501 A -> E (IN REF. 1).
 SQ SEQUENCE 759 AA: 86033 MW: DBA75CC37B1F74D C9C64;
 Query Match 11.0% Score 80; DB 1; Length 759;
 Best Local Similarity 28.7%; Pred. No. 4.8;
 Matches 23; Conservative 13; Mismatches 15; Indels 34; Gaps 7;
 QY 13 GEPAPMTTADGDEGGMCLSSQGLDWLTAHG-TNVS-----SFSWNG-MA 60
 DB 188 GEGP-----NDGCGD-----MEKGR-DWLHMGATNTSMGKMLSYLGAPEMGNP 235
 QY 61 LNPENYLSKDVGTGATKYYKYYVPVNDG 87
 DB 236 LPPELWLT-----PYFLDHPG 252

RESULT 4
 THER_BACTH STANDARD: PRT: 316 AA.
 AC P00800:
 DT 21-JUL-1986 (rel. 01, Created)
 DT 21-JUL-1986 (rel. 01, Last sequence update)
 DT 01-NOV-1997 (rel. 35, Last annotation update)
 DE Thermolysin (EC 3.4.24.27).
 OS Bacillus thermoproteolyticus.
 OC Bacillia; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1427;
 RN
 RP
 RA Tiliant K., Hermodson M.A., Ericsson L.H., Walsh K.A., Neurath H.;
 RT "Amino-acid sequence of thermolysin."
 RL Nature New Biol. 238:35-37(1972).
 RN
 RP
 RA Bursstein Y., Walsh K.A., Neurath H.;
 RT "Evidence of an essential histidine residue in thermolysin."
 RL Biochemistry 13:205-210(1974).
 RN
 RP
 RA X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS).
 RX MEDLINE=83085812; PubMed=7175940;
 RA Holmes M.A., Matthews B.W.;
 RT "Structure of thermolysin refined at 1.6-A resolution."
 RL J. Mol. Biol. 160:623-639(1982).
 RN
 RP
 RA X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
 RX MEDLINE=75041142; PubMed=4214815;
 RA Matthews B.W., Weaver L.H., Kester W.R.;
 RT "The conformation of thermolysin."
 RL J. Biol. Chem. 249:8030-8044(1974).
 RN
 RP
 RA STRUCTURE BY NMR OF 255-316.
 RX MEDLINE=95086079; PubMed=7993910;
 RA Rico M., Jimenez M.A., Gonzalez C., de Filippis V., Fontana A.;
 RT "NMR solution structure of the C-terminal fragment 255-316 of thermolysin: a dimer formed by subunits having the native structure."
 RL Biochemistry 33:14834-14847(1994).
 RN
 RP
 RA STRUCTURE BY NMR OF 205-316.
 RX MEDLINE=97452621; PubMed=9305992;
 RA Conejero-Lara F., Gonzalez C., Jimenez M.A., Padmanabhan S., Maceo P.L., Rico M.;
 RT "NMR solution structure of the 205-316 C-terminal fragment of thermolysin. An example of dimerization coupled to partial unfolding."
 RL Biochemistry 36:11975-11983(1997).
 CC
 CC -1- FUNCTION: THERMOSTABLE EXTRACELLULAR ZINC METALLOPROTEASE.
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Xaa-I-Lau > Xaa-I-Phe.
 CC -1- COFACTOR: BINDS 1 ZINC ION AND FOUR CALCIUM IONS WHICH CONTRIBUTE TO THE UNUSUAL THERMAL STABILITY OF THIS ENDOPEPTIDASE.
 CC TWO ARE CLOSE TOGETHER AND PARTICIPATE IN A NETWORK OF SALT LINKS ALSO INVOLVING ASP-138, GLU-177, LYS-182, ASP-185, GLU-190, AND

CC ASP-191. A THIRD INTERACTS WITH ASP-57 AND A FOURTH WITH ASP-200.
 CC GLU AT POSITION 302 APPEARS TO FORM A SALT LINK WITH LYS-262.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M4.
 DR PIR: A00993; HYBST.
 DR PDB: 4TLN; 31-JAN-84.
 DR PDB: 5TLN; 31-JAN-84.
 DR PDB: 7TLN; 31-JAN-84.
 DR PDB: 8TLN; 30-APR-94.
 DR PDB: 17MN; 09-JAN-89.
 DR PDB: 27MN; 15-JAN-91.
 DR PDB: 37MN; 09-JAN-89.
 DR PDB: 47MN; 09-JAN-89.
 DR PDB: 57MN; 09-JAN-89.
 DR PDB: 67MN; 09-JAN-89.
 DR PDB: 77MN; 15-OCT-94.
 DR PDB: 1THL; 31-JAN-94.
 DR PDB: 1TLF; 09-JAN-89.
 DR PDB: 1HVT; 31-JUL-94.
 DR PDB: 1TRL; 07-FEB-95.
 DR PDB: 1LNA; 08-MAY-95.
 DR PDB: 1LNB; 08-MAY-95.
 DR PDB: 1LNC; 08-MAY-95.
 DR PDB: 1LND; 08-MAY-95.
 DR PDB: 1LNE; 08-MAY-95.
 DR PDB: 1LNF; 08-MAY-95.
 DR MEROPS; M04.001; -;
 DR InterPro: IPR001570; Peptidase_M4.
 DR InterPro: IPR000130; Zn_MTPeptide.
 DR Pfam: PF01447; Peptidase_M4_1.
 DR Pfam: PF02868; Peptidase_M4_C_1.
 DR PRINTS: PR00730; THERMOLYSIN.
 DR PROSITE: PS00142; ZINC_PROTEASE; 1.
 DR HydroLase: Metalloprotease; Zinc.
 KM METAL 142 142
 FT ACT_SITE 143 143
 FT METAL 146 146
 FT METAL 166 166
 FT ACT_SITE 231 231
 FT STRAND 4 11
 FT TURN 13 14
 FT STRAND 17 24
 FT STRAND 28 29
 FT STRAND 31 32
 FT TURN 36 37
 FT STRAND 39 43
 FT TURN 45 46
 FT STRAND 53 54
 FT STRAND 56 57
 FT STRAND 61 62
 FT HELIX 65 88
 FT TURN 92 94
 FT TURN 100 106
 FT TURN 107 109
 FT STRAND 113 113
 FT STRAND 120 123
 FT STRAND 130 130
 FT HELIX 133 135
 FT HELIX 137 151
 FT TURN 152 153
 FT HELIX 159 179
 FT TURN 180 181
 FT STRAND 187 188
 FT TURN 192 192
 FT STRAND 193 193
 FT TURN 195 196
 FT TURN 198 199
 FT TURN 203 204
 FT TURN 212 211
 FT TURN 212 212
 FT HELIX 215 219
 FT HELIX 225 229
 FT TURN 230 233

```

CC      HELIX      234      246
CC      STRAND      248      250
CC      TURN      251      252
CC      STRAND      253      255
CC      HELIX      260      273
CC      TURN      274      274
CC      TURN      277      278
CC      HELIX      281      296
CC      TURN      298      299
CC      HELIX      301      312
CC      TURN      313      314
CC      SEQUENCE 316 AA; 004EDAF478744BB4 CRC64;

Query Match      10.9%; Score 79.5; DB 1; Length 316;
Best Local Similarity 24.7%; Pred. No. 2;
Matches 40; Conservative 28; Mismatches 53; Indels 41; Gaps 10;

OY      9 SSTRH--GEAPAEWT--TIDADGDEGMCLSSGOLDMLTAGGTNNVSSFSNMGMLN 62
DB      102 SSVHYSGYNNAPFNGSOMVYGGDGGDTFIPLSGC-ID-VVAHELTHAVTDYT-AGLIYQ 158
OY      63 PDNYLISK--DYTGATKVKYYPVNDGF-----PG-----DHY 93
DB      159 NESGAINEAISDIFG-TLVEFYANKNPMEIGEEDVYTPGISGDSLRSMSDPAKYGDPDHY 217
OY      94 AVMISKGTGNAGDF--TVVEEFTPNGINKGARGGLSTEANG 133
DB      218 SKRYTGTQDNGGVHINSGIINKAAVLISGCGTHGVSVYIG 259

RESULT 5
NPRS.BACST
ID      NPRS.BACST      STANDARD;      PRT;      551 AA.
AC      P43133;
DT      01-NOV-1995 (Rel. 32, Created)
DT      01-NOV-1995 (Rel. 32, Last sequence update)
DT      01-OCT-1996 (Rel. 34, Last annotation update)
DE      Bacillolysin precursor (EC 3.4.24.28) (Neutral protease).
GN      NPRS OR NPRM.
OS      Bacillus stearothermophilus.
OC      Bacteria; Firmicutes; Bacillales; Geobacillius.
ON      NCBI_TaxID=1422;
RX      MEDLINE=90368536; PubMed=2203733;
RA      Mishima Y., Imanaka T.;
RT      "Cloning and nucleotide sequences of the Bacillus stearothermophilus
RT      neutral protease gene and its transcriptional activator gene.";
RL      J. Bacteriol. 172:4861-4869(1990).
RN      [2]
RP      SEQUENCE FROM N.A., AND SEQUENCE OF 236-240.
RC      STRAIN=MK232;
RA      Kubo N., Imanaka T.;
RT      "Cloning and nucleotide sequence of the highly thermostable neutral
RT      protease gene from Bacillus stearothermophilus.";
RL      J. Gen. Microbiol. 134:1883-1892(1988).
CC      -1- FUNCTION: THERMOSTABLE EXTRACELLULAR ZINC METALLOPROTEASE.
CC      -1- CATALYTIC ACTIVITY: Similar, but not identical, to that of
CC      thermolysin.
CC      -1- COFACTOR: BINDS 1 ZINC ION AND FOUR CALCIUM IONS (BY
CC      SIMILARITY).
CC      -1- SUBCELLULAR LOCATION: Secreted.
CC      -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M4.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).

```

```

CC      EMBL: M34237; AAA2625.1; -
CC      DR EMBL: M21663; AAB02774.1; -
CC      DR HSSP: P00800; 1TR.L.
CC      DR MEROPS: M04.018; -
CC      DR InterPro: IPR005075; pep_M4_propep.
CC      DR InterPro: IPR001570; peptidase_M4.
CC      DR InterPro: IPR000130; zn_mpeptide.
CC      DR pfam: PF01447; peptidase_M4.1.
CC      DR pfam: PF02868; peptidase_M4.C.1.
CC      DR pfam: PF03813; pep_M4_propep.1.
CC      DR PRINTS: PR00730; THERMOLYSIN.
CC      DR PROSITE: PS00142; ZINC_PROTEASE.1.
CC      DR HydroLase; Metalloprotease; zinc: Calcium; zymogen; signal.
CC      FT SIGNAL      1      31
CC      FT PROPEP      32      235
CC      FT CHAIN      236      551
CC      FT METAL      377      377
CC      FT ACT_SITE      378      378
CC      FT METAL      381      381
CC      FT METAL      401      401
CC      FT ACT_SITE      466      466
CC      FT CONFLICT      10      14
CC      FT CONFLICT      24      24
CC      FT CONFLICT      29      33
CC      FT CONFLICT      114      114
CC      FT CONFLICT      124      124
CC      FT CONFLICT      134      134
CC      FT CONFLICT      261      261
CC      FT CONFLICT      463      463
CC      FT CONFLICT      463      463
CC      SO SEQUENCE 551 AA; 60616 MW; FCF4B25A7870129 CRC64;

Query Match      10.9%; Score 79.5; DB 1; Length 551;
Best Local Similarity 24.7%; Pred. No. 3.7;
Matches 40; Conservative 28; Mismatches 53; Indels 41; Gaps 10;

OY      9 SSTRH--GEAPAEWT--TIDADGDEGMCLSSGOLDMLTAGGTNNVSSFSNMGMLN 62
DB      337 SSVHYSGYNNAPFNGSOMVYGGDGGDTFIPLSGC-ID-VVAHELTHAVTDYT-AGLIYQ 393
OY      63 PDNYLISK--DYTGATKVKYYPVNDGF-----PG-----DHY 93
DB      394 NESGAINEAISDIFG-TLVEFYANKNPMEIGEEDVYTPGISGDSLRSMSDPAKYGDPDHY 452
OY      94 AVMISKGTGNAGDF--TVVEEFTPNGINKGARGGLSTEANG 133
DB      453 SKRYTGTQDNGGVHINSGIINKAAVLISGCGTHGVSVYIG 494

RESULT 6
GALM.STRTT
ID      GALM.STRTT      STANDARD;      PRT;      348 AA.
AC      P21953;
DT      01-AUG-1991 (Rel. 19, Created)
DT      01-AUG-1991 (Rel. 19, Last sequence update)
DT      15-JUN-2002 (Rel. 41, Last annotation update)
DE      Aldose 1-epimerase (EC 5.1.3.3) (Mutarotase).
GN      GALM.
OS      Streptococcus thermophilus.
OC      Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC      Streptococcus.
CX      NCBI_TaxID=1308;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=A147;
RX      MEDLINE=90298833; PubMed=1694527;
RA      Poolman B., Royer T.J., Maizner S.E., Schmidt B.F.;
RT      "Carbohydrate utilization in Streptococcus thermophilus:
RT      characterization of the genes for aldose 1-epimerase (mutarotase) and
RT      UDP-glucose 4-epimerase.";
RL      J. Bacteriol. 172:4037-4047(1990).
CC      -1- FUNCTION: Mutarotase converts alpha-aldose to the beta-anomer. It
CC      is active on D-glucose, L-arabinose, D-xylose, D-galactose,

```

CC -1- maltose and lactose (By similarity).
 CC -1- CATALYTIC ACTIVITY: Alpha-D-glucose -> beta-D-glucose.
 CC -1- PATHWAY: Hexose metabolism.
 CC -1- SIMILARITY: BELONGS TO THE ALDOSE EPIMERASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.ebi.ac.uk/announcements/>
 CC or send an email to license@ebi.ac.uk).
 CC
 DR EMBL: M38175; AAA26945.1; -
 DR PIR: B44509; B44509.
 DR InterPro: IPR001823; Ald1-epimerase.
 DR Pfam: PF01263; Aldose-epim: 1.
 DR ProDom: PD005407; Ald1-epimerase: 1.
 DR PROSITE: PS00545; ALDOSE_1-EPIMERASE: 1.
 DR Isoemerase.
 FT ACT_SITE 180 180 POTENTIAL.
 SQ SEQUENCE 348 AA; 39012 MW; 17C3E027ABE1F92 CRC64;
 Query Match 10.8%; Score 78.5; DB 1; length 348;
 Best Local Similarity 31.0%; Pred. No. 2.8; Mismatches 17; Gaps 3;
 Matches 26; Conservative 10; Indels 31; Gaps 3;
 Oy 29 EGMGLSSGQDMHTAGGTNVSSFSNMGMLNPNTLISKDVTGATKKYYPVNDGFP 88
 Db 99 EGDNCL-----HGSPKMGVQNNVYTNLDNDYETK-----FIRRLYSVDGF 142
 Oy 89 PCDHAYVMISYGTNGCDFTVVE 112
 Db 143 PGD-VTVSISYRLNNRRILTFE 165
 RESULT 7
 TNA1_SYMTM
 ID TNA1_SYMTM STANDARD; PRT: 453 AA.
 AC P31013;
 DT 01-JUN-1993 (Rel. 26, Created)
 DT 01-JUN-1993 (Rel. 26, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Tryptophanase 1 (EC 4.1.99.1) (L-tryptophan indole-lyase 1) (Tnase 1).
 GN TNA1
 OS Symbiobacterium thermophilum.
 OC Bacteria; Actinobacteria; Actinobacteria (class); Symbiobacterium.
 OX NCBI_TaxID:2734;
 RN [1]
 RP SEQUENCE FROM N.A. AND SEQUENCE OF 1-36.
 RC STRAIN: IAM_115621; PubMed:1339259;
 RX MEDLINE:92384583;
 RA Hirshara T., Suzuki S., Morinouchi S., Beppu T.;
 "Cloning, nucleotide sequences, and overexpression in *Escherichia*
 coli of tandem copies of a tryptophanase gene in an obligately
 KT symbiotic thermophilic Symbiobacterium thermophilum.";
 RL Appl Environ Microbiol. 58:2633-2642(1992).
 CC -1- CATALYTIC ACTIVITY: L-tryptophan + H(2)O = indole + pyruvate +
 NH(3)
 CC -1- COFACTOR: pyridoxal phosphate.
 CC -1- PATHWAY: tryptophan catabolism.
 CC -1- SUBUNIT: homodimer (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE BETA-ELIMINATING LYASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.ebi.ac.uk/announcements/>
 CC or send an email to license@ebi.ac.uk).
 CC
 DR EMBL: AB010632; BAA24668.1; -

DR PIR: A49022; A49022.
 DR HSSP: P31013; 1TPU.
 DR InterPro: IPR001597; Beta_elim_lyase.
 DR Pfam: PF01212; Beta_elim_lyase: 1.
 DR ProDom: PD005927; Beta_elim_lyase: 1.
 DR PROSITE: PS00853; BETA_ELIM_LYASE: 1.
 DR Tryptophan catabolism: Lyase; Pyridoxal phosphate.
 FT BINDING 254 254 PYRIDOXAL PHOSPHATE (By similarity).
 SQ SEQUENCE 453 AA; 51012 MW; CAB332655C08812 CRC64;
 Query Match 10.8%; Score 78.5; DB 1; length 453;
 Best Local Similarity 26.6%; Pred. No. 3.7; Mismatches 41; Indels 13; Gaps 4;
 Matches 25; Conservative 15; Mismatches 41; Indels 13; Gaps 4;
 Oy 39 LDMHTAGGTNVSSFSNMGMLNPNTL-----ISKDVTGATKKYYPVNDGFP 91
 Db 44 IDLLT-DSGTGMSGFQSMMLGDESYAGASVYRLKERTVDITGYVPIPHGRGAE 102
 Oy 92 HYAMISKGTAGDF--TVFEETPNCINKG 122
 Db 103 --KVAFSQILTPGMYVLSMFPDTRGHVQLG 134
 RESULT 8
 GUNZ_ERWCH
 ID GUNZ_ERWCH STANDARD; PRT: 426 AA.
 AC P07103;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Endoglucanase 2 precursor (EC 3.2.1.4) (Endo-1,4-beta-glucanase 2)
 DE (Cellulase 2) (EC2).
 GN CE1Z OR CE1S.
 OS *Erwinia chrysanthemi*.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Pectobacterium.
 OX NCBI_TaxID:556;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN: 3937.
 RX MEDLINE:86216177; PubMed:2835589;
 RA Guseppl A., Caml B., Kymecic J.-L., Ball G., Creuset N.;
 "Homology between endoglucanase 2 of *Erwinia chrysanthemi* and
 RT endoglucanase of *Bacillus subtilis* and alkalophilic *Bacillus*.";
 RU Mol. Microbiol. 2:159-164(1988).
 RN [2]
 RP REVISIONS: AND DISULFIDE BOND.
 RC STRAIN: 3937.
 RX MEDLINE:94203057; PubMed:6152378;
 RA Bortoli-German I., Bryn E., Fry B., Chippaux M., Barras F.;
 RT Periplasmic disulphide bond formation is essential for cellulase
 RT secretion by the plant pathogen *Erwinia chrysanthemi*.";
 RU Mol. Microbiol. 11:545-553(1994).
 RN [3]
 RP MOTIVEMESIS AND DOMAINS.
 RX MEDLINE:91312860; PubMed:1677466;
 RA Fry B., Bortoli-German I., Halech J., Chippaux M., Barras F.;
 RT Cellulase Ec2 of *Erwinia chrysanthemi*: structural organization and
 RT importance of His98 and Glu133 residues for catalysis.";
 RU Protein Eng. 4:325-333(1991).
 RN [4]
 RP STEREOCHEMISTRY OF THE REACTION.
 RX MEDLINE:9222124; PubMed:1563515;
 RA Heyraud A., Bortoli-German I., Bauzan M., Kouvler J., Gey C.,
 RT Stereochemistry of the hydrolysis reaction catalyzed by
 RT endoglucanase 2 from *Erwinia chrysanthemi*.";
 RU FEBS Lett. 300:145-148(1992).
 RN [5]
 RP STRUCTURE BY NMR OF 365-426.
 RX MEDLINE:96070232; PubMed:9405041;
 RA Brun E., Morlaud F., Gans P., Blackledge M.J., Barras F., Marion D.;
 RT "Solution structure of the cellulose-binding domain of the

[illegible]

```

01-NOV-1995 (Rel. 32, Last annotation update)
DB Bacillolysins Precursor (Ec 3.4.24.28) (Neutral protease).
CM NMR.
OS Bacillus brevis.
SC Bacteria; Firmicutes; Bacillales; Paenibacillaceae; Brevibacillus.
OK NCBI_TaxID=1393;
RM (1)
RP SEQUENCE FROM N.A.
RC STRAIN=7882;
RA MEDLINE=91148616; PubMed=2290427;
RA ARAKOV A.S., BOLOTIN A.P., SOROKIN A.V.;
RT "Structure of the Bacillus brevis metalloprotease gene.";
RL Mol. Biol. (Mosk) 24:1363-1372(1990).
RM (2)
RP SEQUENCE OF 224-228, AND CHARACTERIZATION.
RA MEDLINE=91148618; PubMed=2127074;
RA KRIDLOVA N.V., AKIMINA T.V., KHODOVA O.D., KOSTROV S.V.,
RM "Analysis of the structure of Bacillus brevis neutral proteinase and its biosynthesis in Bacillus subtilis cells.";
RT Mol. Biol. (Mosk) 24:1381-1392(1990).
CC - FUNCTION: THERMOSTABLE EXTRACELLULAR ZINC METALLOPROTEASE.
CC - CATALYTIC ACTIVITY: Similar, but not identical, to that of thermolysin.
CC - COFACTOR: BINDS 1 ZINC ION AND FOUR CALCIUM IONS (BY SIMILARITY).
CC - SUBCELLULAR LOCATION: Secreted.
CC - SIMILARITY: BELONGS TO PEPTIDASE FAMILY M4.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X61286; CAA3589.1; -.
DR HSSP; P00800; IHYH.
DR MEROPS; M04_UPM; -.
DR InterPro; IPR005075; Pep_M4_Proprep.
DR InterPro; IPR001570; Peptidase_M4.
DR InterPro; IPR000130; Zn_MTPeptide.
DR Pfam; PF01447; Peptidase_M4_1.
DR Pfam; PF02868; Peptidase_M4_C_1.
DR Pfam; PF03413; Pep_M4_Proprep; 1.
DR PRINTS; PR00730; THERMOLYSIN.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Hydrolyse; Metalloprotease; Zinc; Calcium; Zymogen; Signal.
FT SIGNAL 1 28
FT PROPEP 29 223 POTENTIAL ACTIVATION PEPTIDE.
FT CHAIN 224 527 BACILLOLYSIN.
FT METAL 358 358 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 359 359 BY SIMILARITY.
FT METAL 362 362 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 382 382 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 442 442 PHOTON DONOR (BY SIMILARITY).
SQ SEQUENCE 527 AA; 58646 MW; 8D3704C3CB9D8756 CRC64;
Query Match Best Local Similarity 10.4%; Score 76; DB 1; Length 527;
Matches 37; Conservative 20; Mismatches 43; Indels 40; Caps 8;
0Y 25 DGDGEWMLTSSGGLJMLAHNGCTNVS-----SFSNN-----GNALPNPY 66
Db 340 DCGGTFFLPISGG-LD-VVAHELTHAETERTGLQVNESGALNMSDSIGAWVIDDM 397
0Y 67 LISKRD-----TGATKKYKYYPVDNGFPGGHAVMISTGNGDPVTVAEPENGIN-- 119
Db 398 LMEDIEDIYPPGSGCALRSIQDPAAYGDP-DHYSKRYTSQSDNG-----VHTNSGINNK 450
0Y 120 -----KGARFGSTLEANG 133

```



```
FT DISULFID 126 146 BY SIMILARITY.
FT DISULFID 160 173 BY SIMILARITY.
FT DISULFID 479 488 BY SIMILARITY.
FT DISULFID 494 551 BY SIMILARITY.
FT DISULFID 612 618 BY SIMILARITY.
FT DISULFID 684 697 BY SIMILARITY.
FT DISULFID 838 890 INTERCHAIN (BY SIMILARITY).
FT DISULFID 895 900 BY SIMILARITY.
FT CARBOHYD 62 62 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 278 278 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 284 284 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 540 540 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 601 601 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 690 690 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 821 821 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 837 837 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 860 860 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 931 931 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 951 951 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 959 959 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 966 966 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1034 AA; 114388 MW; D76B0B8A692DC684 CRC64;

Query Match 10.3%; Score 75.5; DB 1; Length 1034;
Best Local Similarity 23.8%; Pred. No. 18;
Matches 36; Conservative 21; Mismatches 47; Indels 47; Gaps 9;

OY 7 FESSTHGE-ANAEMTTTADGDGEGMLCLSSGQDLMLTAHGNTNV-----SFSNMGMAL 61
DB 149 FDGSKSVYAPACRSYTTIDADGCG---FCGGFSIDFTK---GDRVLLGSGPSFYWGOL 202
OY 62 NP-----DNVLISKDYTGATKYKYYPVNDGFPFGDHYAVAMISKTGTN 103
DB 203 SDRAVELAKKDSKYSTKYIDQLATPRASAA-----FDDSYLG--YSVAV----- 246
OY 104 AGDFT-VVEETPNGINKGARFGLSTENG 133
DB 247 -GDFSGDIEDFVSGVPRARLGMVSIYNG 276

RESULT 12
THER_BACST
ID THER_BACST STANDARD: PRT; 548 AA.
AC P06874;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Thermolysin precursor (EC 3.4.24.27) (Thermostable neutral
  protease).
GN NPRT.
OS Bacillus stearothermophilus.
OC Bacteria; Firmicutes; Bacillales; Geobacillus.
OX NCBI_TaxID=1422;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 230-243.
RX MEDLINE=85289021; PubMed=2933245;
RA Takegi M., Imanaka T., Alba S.;
RT "Nucleotide sequence and promoter region for the neutral protease
  gene from Bacillus stearothermophilus.";
RL J. Bacteriol. 163:824-831(1985).
CC -1- FUNCTION: THERMOSTABLE EXTRACELLULAR ZINC METALLOPROTEASE.
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Xaa-I-Leu > Xaa-I-Phe.
CC -1- COFACTOR: BINDS 1 ZINC ION AND FOUR CALCIUM IONS.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M4.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
CC -----
DR EMBL: M11446; AAA22621.1;
DR PIR: A24824; HYBS.
DR HSSP: P00800; IHVT.
DR MEROPS: M04.001; -.
DR InterPro: IPR005075; Pep_M4_propep.
DR InterPro: IPR001570; Peptidase_M4.
DR InterPro: IPR000130; Zn_mtpeptide.
DR Pfam: PF01447; Peptidase_M4_1.
DR Pfam: PF02866; Peptidase_M4_C1.
DR Pfam: PF03413; Pep_M4_propep; 1.
DR PRINTS: PR00730; THERMOLYSIN.
DR PROSITE: PS00142; ZINC_PROTEASE; 1.
KW Hydrolyase; Metalloprotease; Zinc; Calcium; Zymogen; Signal.
FT SIGNAL 1 25
FT PROPEP 26 229 ACTIVATION PEPTIDE.
FT CHAIN 230 548 THERMOLYSIN.
FT METAL 374 374 ZINC (CATALYTIC).
FT ACT_SITE 375 375
FT METAL 378 378 ZINC (CATALYTIC).
FT METAL 398 398 ZINC (CATALYTIC).
FT ACT_SITE 463 463 PROTON DONOR.
SQ SEQUENCE 548 AA; 59580 MW; 584D2C1D7910539D CRC64;

Query Match 10.2%; Score 74.5; DB 1; Length 548;
Best Local Similarity 23.5%; Pred. No. 11;
Matches 38; Conservative 28; Mismatches 55; Indels 41; Gaps 10;

OY 9 GSTH---GEAPAEWT---TIDADGDEGMLCLSSGQDLMLTAHGNTNVSSFSNMGMALN 62
DB 334 STVHYRGYNNAFNGSQMYVGDGQTFLEPFGS-G-ID-VVGHDLTHAVDTY-AGLVYQ 390
OY 63 PDNLISK---DVGATKVKYKYYPVNDGF-----DG-----DHY 93
DB 391 NESGAINEAMSDIFG-TLVEFYANRPMEIGEDITYTGAVAGDALRMSDPAKYGDPRHY 449
OY 94 AVMISKGTGTNAGDFT-VVEETPNGINKGARFGLSTENG 133
DB 450 SKRYTGQDNGCVHTNSGIINKAAVLLSQGVHYGVSVNGIG 491

RESULT 13
TNA2_SYMTH
ID TNA2_SYMTH STANDARD: PRT; 452 AA.
AC P31015; G9R5M8;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tryptophanase 2 (EC 4.1.99.1) (L-tryptophan indole-lyase 2) (Tnase 2).
GN TNA2.
OS Symbiodacterium thermophilum.
OC Bacteria; Actinobacteria; Actinobacteria (class); Symbiodacterium.
OX NCBI_TaxID=2734;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-34.
RX STRAIN=IAM 13621;
RX MEDLINE=92384583; PubMed=1339259;
RA Hirahara T., Suzuki S., Horinouchi S., Beppu T.;
RT "Cloning, nucleotide sequences, and overexpression in Escherichia
  coli of tandem copies of a tryptophanase gene in an obligately
  symbiotic thermophile, Symbiodacterium thermophilum.";
RL Appl. Environ. Microbiol. 58:2633-2642(1992).
RN [2]
RP SEQUENCE OF 1-35.
RX STRAIN=T;
RX MEDLINE=92144113; PubMed=1368766;
RX Suzuki S., Hirahara T., Horinouchi S., Beppu T.;
RT "Purification and properties of thermostable tryptophanase from an
  obligately symbiotic thermophile, Symbiodacterium thermophilum.";
RL Agric. Biol. Chem. 55:3059-3066(1991).
CC -1- CATALYTIC ACTIVITY: L-tryptophan + H(2)O = indole + pyruvate +
CC NH(3).
CC -1- COFACTOR: Pyridoxal phosphate.
```

```
CC -1- PATHWAY: Tryptophan catabolism.
CC -1- SUBUNIT: Homotetramer.
CC -1- SIMILARITY: BELONGS TO THE BETA-ELIMINATING LYASE FAMILY.
CC -1- CAUTION: TNA2 IS EXPRESSED IN STRAIN T BUT EXPRESSION IN STRAIN
CC IAM 13621 HAS NOT BEEN PROVEN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: AB010832; BAA24687.1; ..
CC PIR: B49022; B49022.
CC HSSP: P31013; 1PFL.
CC InterPro: IPR001597; Beta_elim_lyase.
CC Pfam: PF01212; Beta_elim_lyase.1.
CC ProDom: PD005927; Beta_elim_lyase.1.
CC PROSITE: PS00853; BETA_ELIM_LYASE.1.
CC Tryptophan catabolism; Lyase; Pyridoxal phosphate.
CC INIT_MET 0
CC BINDING 253 253 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
CC SEQUENCE 452 AA; 50472 MW; C442971C2CA43138 CRC64;
CC -----
Query Match 10.1%; Score 74; DB 1; Length 452;
Best Local Similarity 27.8%; Pred. No. 9.9;
Matches 20; Conservative 14; Mismatches 28; Indels 10; Gaps 3;
CC -----
OY 39 LDMLTANGCTNVSSFSNGMALNPNDNYL-----ISKDYTGATKVKYYPVNDGPPG 91
DB 43 IDLLT-DSGTGAMSGFOWSAMLDCDSYACASRYLKEAVTDITGEYVLPHQGCAE 101
OY 92 H-YAVMISKTG 101
DB 102 KSAFAQLITRPG 113
CC -----
RESULT 14
GNNW_ERMCA STANDARD; PRT: 504 AA.
ID AC 053935;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Endoglucanase VI precursor (EC 3.2.1.4) (Endo-1,4-beta-glucanase VI)
DE (Cellulase VI).
CN CELV1.
OS Erythrina carotovora.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Pectobacterium.
OX NCBI_TaxID=554;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SCC3193;
RC MEDLINE=95231512; PubMed=7715600;
RC Mae A., Heikinheimo R., Palva E.T.;
RT "Structure and regulation of the Erythrina carotovora subspecies
RT carotovora SCC3193 cellulase gene celv1 and the role of cellulase in
RT phytopathogenicity.";
RT Mol. Gen. Genet. 247:17-26(1995).
RL -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic
CC linkages in cellulose.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO CELLULOSE FAMILY A (FAMILY 5 OF GLYCOSYL
CC HYDROLASES).
CC -----
This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
```

```
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: X79241; CA55823.1; ..
CC HSSP: O85465; 1A3H.
CC InterPro: IPR001956; CBD_3.
CC InterPro: IPR001547; GH_5.
CC Pfam: PF00150; cellulase.1.
CC Pfam: PF00942; CBM_3; 1.
CC ProDom: PD001947; CBD_3; 1.
CC PROSITE: PS00659; GLYCOSYL_HYDROL_F5; 1.
CC Cellulose degradation; Hydrolase; Glycosidase; Signal.
CC SIGNAL 1
CC CHAIN 31
CC FT CHAIN 32 504
CC FT DOMAIN 32 334
CC FT DOMAIN 335 352
CC FT DOMAIN 353 504
CC FT ACT_SITE 168
CC FT ACT_SITE 256 256
CC SEQUENCE 504 AA; 54963 MW; 0D7ECF74781565FA CRC64;
CC -----
Query Match 10.1%; Score 74; DB 1; Length 504;
Best Local Similarity 20.7%; Pred. No. 11;
Matches 34; Conservative 21; Mismatches 65; Indels 44; Gaps 5;
CC -----
OY 7 FEESTHCE-----APAETTDADGEGWCLSGQDLMTANGTNV 50
DB 229 FYGTGHEQFLRDRIIDYAQSGAIIYSEWGTSDASGNGGPFPEQSTWIDFLNNNG--- 284
OY 51 VSSFSNGMALNPNDNYLSKDYTGATK-----VKYYPVNDGPPG 95
DB 285 ---ISWVNMSLSDKSEFSAALVAGASKSGGWFGQNLSTGSKFVREQIRAGAGLSGDDPT 341
OY 96 MISK-----TGTNAGDEPTVFEETPNKIGKGRFELSPENCA 134
DB 342 MPTPTNPGNGT-TGDVILQYRVNDNPNPSDDAIRMAFNKNTGS 384
CC -----
RESULT 15
THT2_MYCTU STANDARD; PRT: 297 AA.
ID AC P66888;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Putative thiosulfate sulfurtransferase SSEA (EC 2.8.1.1).
DE SSEA OR RV3283 OR MT3382 OR MTCY71.23.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37Rv;
RC MEDLINE=98295987; PubMed=9634230;
RC Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RC Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
RC Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RC Davies R., Devlin K., Feltham D., Gentles S., Hamlin N., Holtroyd S.,
RC Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RC Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RC Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RC Stalston J.E., Taylor K., Whitehead S., Barrett B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RT Nature 393:537-544(1998).
RL [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RC Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RC Peterson J., Deboy R., Dodson R., Gwin M.L., Haft D., Hickey E.,
RC Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RC Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
```


Thu Jun 26 11:58:17 2003

us-09-980-370-6.rsp

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 26, 2003, 00:47:39 ; Search time 40.9882 Seconds
(without alignments)
673.616 Million cell updates/sec

Title: US-09-980-370-6

Perfect score: 730
Sequence: 1 ADFTETFEESSTHGEPAEMWT.....PNCINKGCAFGSLTEANCA 134

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

1: SP-archaea:.*
2: SP-bacteria:.*
3: SP-fungi:.*
4: SP-human:.*
5: SP-invertebrate:.*
6: SP-mammal:.*
7: SP-misc:.*
8: SP-organella:.*
9: SP-phage:.*
10: SP-plant:.*
11: SP-rodent:.*
12: SP-virus:.*
13: SP-vertebrate:.*
14: SP-unclassified:.*
15: SP-virus:.*
16: SP-bacteriophage:.*
17: SP-archaeop:.*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Match	Length	DB	ID	Description
1	715	98.1	1223	2	092NB5	092NB5 porphyromon
2	715	98.1	1687	2	09R8B7	09R8B7 porphyromon
3	715	98.1	1704	2	051816	051816 porphyromon
4	715	97.9	1097	2	P72196	P72196 porphyromon
5	715	97.9	1732	2	052050	052050 porphyromon
6	715	97.9	1732	2	051817	051817 porphyromon
7	715	97.9	1732	2	007442	007442 porphyromon
8	714	97.8	1706	2	051839	051839 porphyromon
9	711	97.4	1723	2	P72197	P72197 porphyromon
10	709	97.1	1358	2	P96967	P96967 porphyromon
11	709	97.1	1723	2	P72194	P72194 porphyromon
12	706	96.7	1706	2	051838	051838 porphyromon
13	205	28.1	925	2	09FA40	09FA40 porphyromon
14	180	24.7	312	2	09K1B3	09K1B3 porphyromon
15	90.5	12.4	218	16	09J215	09J215 neisseria m
16	90.5	12.4	286	17	09V211	09V211 pyrococcus

17	89	12.2	418	9	080060	080060 staphylococ
18	89	12.2	1235	9	08SDK2	08SDK2 staphylococ
19	89	12.2	1261	9	09MBN9	09MBN9 staphylococ
20	88.5	12.1	2101	16	09J010	09J010 neisseria m
21	88.5	12.1	1217	16	08Y378	08Y378 neisseria s
22	87	11.9	910	3	P87211	P87211 orpomyces
23	85.5	11.7	2275	17	08TSE8	08TSE8 methanosarc
24	83.5	11.4	256	17	08TSD2	08TSD2 bacillus th
25	83.5	11.4	548	2	045779	045779 bacillus th
26	83.5	11.4	603	10	09M653	09M653 polyonatum
27	83	11.4	373	16	053926	053926 mycobacteri
28	83	11.4	806	3	09P478	09P478 agaricus bi
29	82.5	11.3	536	17	097YR2	097YR2 sulfolobus
30	81.5	11.2	494	2	086039	086039 alteromonas
31	81.5	11.2	966	17	026770	026770 methanobact
32	81	11.1	1372	2	054151	054151 shigella fl
33	81	11.1	1373	2	09AL58	09AL58 shigella fl
34	79.5	10.9	534	5	025403	025403 lymnaea sta
35	79.5	10.9	735	16	08XNF9	08XNF9 clostridium
36	78.5	10.8	348	2	08VS90	08VS90 streptococ
37	78.5	10.8	348	2	08VS84	08VS84 streptococ
38	78	10.7	499	2	045532	045532 bacillus su
39	78	10.7	785	3	09PB85	09PB85 humicola in
40	78	10.7	803	10	08W075	08W075 sorghum bic
41	78	10.7	1506	12	091LD6	091LD6 white spot
42	77.5	10.6	1039	5	09V0B1	09V0B1 dirosophila
43	77	10.5	698	2	09KH03	09KH03 alcaligenes
44	76	10.4	344	16	09BNM4	09BNM4 rhizobium j
45	76	10.4	680	2	052644	052644 rumioccocu

ALIGNMENTS

RESULT 1

ID 092NB5

PRELIMINARY:

PRT: 1223 AA.

AC 092NB5: 1-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE 130K-HMGD (Fragment).
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteri: CPB group: Bacteroidetes; Bacteroidales: Porphyromonadaceae;
OC Porphyromonas.
NCBI_TaxID=837.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=381.
RX MEDLINE=9914316; PubMed=998746;
RA Shibata Y., Hayekawa M., Takiguchi H., Shiroza T., Abiko Y.:
RT "Determination and characterization of the hemagglutinin-associated
RT short motifs found in Porphyromonas gingivalis multiple gene
RT products".
RL J. Biol. Chem. 274:5012-5020(1999).
DR EMBL: AB019363; BAA34341.1; .
DR InterPro: IPR000977; DNA_Ligase.
DR InterPro: IPR001769; Peptidase_C25.
DR Pfam: PF01364; Peptidase_C25; 3.
DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
FT NON_TER
SQ SEQUENCE 1223 AA: 131542 MW: 00225CD2BA9P91B3 CRC64:
Query Match 98.1%; Score 716; DB 2; Length 1223;
Best local similarity 97.8%; Pred. No. 1.2e-54;
Matches 131; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 ADFTETFEESSTHGEPAEMWTIDAGDGGWGLSSGQIDWITAGCTNVSSFSMNCA 60
D6 648 ADFTETFEESSTHGEPAEMWTIDAGDGGWGLSSGQIDWITAGCTNVSSFSMNCA 707
QY 61 LNPENYLSKDVGTGATGATVKKYYPVNDGPGSDHYAVMISGTGNGDPTVFEETPNCINK 120
|||||

```

Db      708 LMPDNYLISKDVTGATKRYKYAVVNDGFGDHVAVMISKTGTNAGDFTVFEETPNCINK 767
Qy      121 GGARFGSLTEANCA 134
        |||
Db      768 GGARFGSLTEANCA 781

RESULT 2
Q9R9B7 PRELIMINARY: PRT: 1687 AA.
ID      Q9R9B7;
DT      01-MAY-2000 (TREMblrel. 13, Created)
DT      01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT      01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE      Hemagglutinin/protease.
GN      HAGE.
OS      Porphyromonas gingivalis (Bacteroides gingivalis).
OC      Bacteriia: CFb group: Bacteroidetes; Bacteroidales; Porphyromonadaceae;
OC      Porphyromonas.
OX      NCBI_TaxID=837;
RN      (1)
RP      SEQUENCE FROM N.A.
RC      STRAIN=381;
RA      Han N., Dong H., Progniske-Fox A.;
RT      "Cloning and characterization of hage from P. gingivalis 381."
RL      Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
DR      EMBL: AF028946; AAD01810.1;
DR      HSSP: P95493; ICVR.
DR      MEROPS: C25.001;
DR      InterPro: IPR000977; DNA_ligase.
DR      InterPro: IPR001769; Peptidase_C25.
DR      Pfam: PF01364; Peptidase_C25; 3.
DR      PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
SQ      SEQUENCE 1687 AA: 183702 MW: D0858516A399FE70 CRC64:

Query Match
Best Local Similarity 98.1%; Score 716; DB 2; Length 1687;
Matches 131; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy      1 ADFETFEESTHGEAPAEWTTIDADGCGWLCSSQGLDMLTAHGTNNVSSFSWNGMA 60
        |||
Db      1120 ADFETFEESTHGEAPAEWTTIDADGCGWLCSSQGLDMLTAHGTNNVSSFSWNGMA 1179
Qy      61 LMPDNYLISKDVTGATKRYKYAVVNDGFGDHVAVMISKTGTNAGDFTVFEETPNCINK 120
        |||
Db      1180 LMPDNYLISKDVTGATKRYKYAVVNDGFGDHVAVMISKTGTNAGDFTVFEETPNCINK 1239
Qy      121 GGARFGSLTEANCA 134
        |||
Db      1240 GGARFGSLTEANCA 1253

RESULT 3
Q51816 PRELIMINARY: PRT: 1704 AA.
ID      Q51816;
DT      01-NOV-1996 (TREMblrel. 01, Created)
DT      01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT      01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE      Arg-gingipain-1 proteinase.
OS      Porphyromonas gingivalis (Bacteroides gingivalis).
OC      Bacteriia: CFb group: Bacteroidetes; Bacteroidales; Porphyromonadaceae;
OC      Porphyromonas.
OX      NCBI_TaxID=837;
RN      (1)
RP      SEQUENCE FROM N.A.
RC      MECLINE-95138080; Pubmed-7636351;
RA      Trevis J., Barr P.J.;
RT      "Molecular cloning and structural characterization of the Arg-
        gingipain proteinase of Porphyromonas gingivalis. Biosynthesis as a
        proteinase-adhesin polypeptide."

```

```

RL      J. Biol. Chem. 270:1007-1010(1995).
DR      EMBL: U15282; AAA69539.1;
DR      HSSP: P95493; ICVR.
DR      MEROPS: C25.001;
DR      InterPro: IPR000977; DNA_ligase.
DR      InterPro: IPR001769; Peptidase_C25.
DR      Pfam: PF01364; Peptidase_C25; 3.
DR      PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
FT      CHAIN 228 MATURE 50-KDA CYSTEINE PROTEINASE
        GINGIPAIN.
SQ      SEQUENCE 1704 AA: 185436 MW: 6A34B40131C2A676 CRC64:

Query Match
Best Local Similarity 98.1%; Score 716; DB 2; Length 1704;
Matches 131; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy      1 ADFETFEESTHGEAPAEWTTIDADGCGWLCSSQGLDMLTAHGTNNVSSFSWNGMA 60
        |||
Db      1137 ADFETFEESTHGEAPAEWTTIDADGCGWLCSSQGLDMLTAHGTNNVSSFSWNGMA 1196
Qy      61 LMPDNYLISKDVTGATKRYKYAVVNDGFGDHVAVMISKTGTNAGDFTVFEETPNCINK 120
        |||
Db      1197 LMPDNYLISKDVTGATKRYKYAVVNDGFGDHVAVMISKTGTNAGDFTVFEETPNCINK 1256
Qy      121 GGARFGSLTEANCA 134
        |||
Db      1257 GGARFGSLTEANCA 1270

RESULT 4
P72196 PRELIMINARY: PRT: 1097 AA.
ID      P72196;
DT      01-FEB-1997 (TREMblrel. 02, Created)
DT      01-FEB-1997 (TREMblrel. 02, Last sequence update)
DT      01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE      Tumb-linked adhesin precursor.
GN      TLA.
OS      Porphyromonas gingivalis (Bacteroides gingivalis).
OC      Bacteriia: CFb group: Bacteroidetes; Bacteroidales; Porphyromonadaceae;
OC      Porphyromonas.
OX      NCBI_TaxID=837;
RN      (1)
RP      SEQUENCE FROM N.A.
RC      MECLINE-97386416; Pubmed-9244265;
RA      Aduse-Opoku J., Stanley J.M., Young K.A., Muir J., Rangarajan M.,
        "TumCis M.A.";
RT      "The tla gene of Porphyromonas gingivalis W50: a homologue of the
        arginine-specific protease precursor (PipRI) which shares sequence
        similarity to 1208-1778."
DR      EMBL: Y07618; GA56897.1;
DR      InterPro: IPR000977; DNA_ligase.
DR      InterPro: IPR001769; Peptidase_C25.
DR      Pfam: PF01364; Peptidase_C25; 3.
DR      PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
FT      SIGNAL 1
SQ      SEQUENCE 1097 AA: 118731 MW: 73BBH3378421F8B9 CRC64:

Query Match
Best Local Similarity 97.8%; Score 715; DB 2; Length 1097;
Matches 131; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy      1 ADFETFEESTHGEAPAEWTTIDADGCGWLCSSQGLDMLTAHGTNNVSSFSWNGMA 60
        |||
Db      522 ADFETFEESTHGEAPAEWTTIDADGCGWLCSSQGLDMLTAHGTNNVSSFSWNGMA 581
Qy      61 LMPDNYLISKDVTGATKRYKYAVVNDGFGDHVAVMISKTGTNAGDFTVFEETPNCINK 120
        |||
Db      582 LMPDNYLISKDVTGATKRYKYAVVNDGFGDHVAVMISKTGTNAGDFTVFEETPNCINK 641
Qy      121 GGARFGSLTEANCA 134

```

Db 642 GGARFGLSTEANGA 655

RESULT 5
ID 052050 PRELIMINARY: PRT: 1732 AA.

AC 052050: 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
OS Lysine specific cysteine protease.
OC Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria: CFB group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
OC Porphyromonas.
OX NCBI_TaxID=837;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-W83:
RX MEDLINE=98298016; PubMed=9632563;
RA Lewis J.P., Macrina F.L.,
RT -IS195, an insertion sequence-like element associated with protease
RT genes in Porphyromonas gingivalis.";
RL Infect. Immun. 66:3035-3042(1998).
DR EMBL: AF017059; AAC26523.1; -.
DR MEROPS: C25.002; -.
DR InterPro: IPR000977; DNA_ligase.
DR InterPro: IPR001769; Peptidase_C25.
DR Pfam: PF01364; Peptidase_C25; 3.
DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
KW Protease.
SQ SEQUENCE 1732 AA: 187931 MW: 823746305CB5EA5 CRC64:

Query Match 97.9%: Score 715; DB 2: Length 1732;
Best Local Similarity 97.8%: Pred. No. 2,4e-54;
Matches 131: Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 ADFTETFEESTHGEPAEMTTIDADGEGMCLSSGQDMLTAHGCTNVVSSFSWNGMA 60
DB 1157 ADFTETFEESTHGEPAEMTTIDADGEGMCLSSGQDMLTAHGCTNVVSSFSWNGMA 1216
OY 61 LNPBNYLISKDVYGTATVKYKYYYPVNDGFPGDHYAVMISKTGTNAGDFTVVEETPNCINK 120
DB 1217 LNPBNYLISKDVYGTATVKYKYYAVNDGFPGDHYAVMISKTGTNAGDFTVVEETPNCINK 1276
OY 121 GGARFGLSTEANGA 134
DB 1277 GGARFGLSTEANGA 1290

RESULT 6
ID 051817 PRELIMINARY: PRT: 1732 AA.
AC 051817:
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
OC Porphyromonas.
OC Bacteria: CFB group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
OC Porphyromonas.
OX NCBI_TaxID=837;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-W42:
RX MEDLINE=96213011; PubMed=8631659;
RA Barkocy-Gallagher G.A., Han N., Patel J.M., Whitlock J.,
RT Prokulske-Fox A., Lantz M.S.;
RT "Analysis of the prtP gene encoding porphyrin, a cysteine proteinase
RT of Porphyromonas gingivalis.";
RL J. Bacteriol. 178:2734-2741(1996).
DR EMBL: U42210; AAB06565.1; -.

DR MEROPS: C25.002; -.
DR InterPro: IPR000977; DNA_ligase.
DR InterPro: IPR001769; Peptidase_C25.
DR Pfam: PF01364; Peptidase_C25; 3.
DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
SQ SEQUENCE 1732 AA: 187875 MW: 654271DBEF7BCA64 CRC64:

Query Match 97.9%: Score 715; DB 2: Length 1732;
Best Local Similarity 97.8%: Pred. No. 2,4e-54;
Matches 131: Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 ADFTETFEESTHGEPAEMTTIDADGEGMCLSSGQDMLTAHGCTNVVSSFSWNGMA 60
DB 1157 ADFTETFEESTHGEPAEMTTIDADGEGMCLSSGQDMLTAHGCTNVVSSFSWNGMA 1216
OY 61 LNPBNYLISKDVYGTATVKYKYYYPVNDGFPGDHYAVMISKTGTNAGDFTVVEETPNCINK 120
DB 1217 LNPBNYLISKDVYGTATVKYKYYAVNDGFPGDHYAVMISKTGTNAGDFTVVEETPNCINK 1276
OY 121 GGARFGLSTEANGA 134
DB 1277 GGARFGLSTEANGA 1290

RESULT 7
ID 007442 PRELIMINARY: PRT: 1732 AA.
AC 007442:
DT 01-JUL-1997 (TREMBlrel. 04, Created)
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
OS Lysine-specific cysteine proteinase.
OC Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria: CFB group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
OC Porphyromonas.
OX NCBI_TaxID=837;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-W50:
RX MEDLINE=99235907; PubMed=10219167;
RA Stakessi N., Cleal S.M., Bhogal P.S., Reynolds E.C.;
RT "Characterization of a porphyromonas gingivalis gene prtK that encodes
RT a lysine-specific cysteine proteinase and three sequence-related
RT adhesins.";
RL Oral Microbiol. Immunol. 14:92-97(1999).
DR EMBL: U75366; AAB60809.1; -.
DR MEROPS: C25.002; -.
DR InterPro: IPR000977; DNA_ligase.
DR InterPro: IPR001769; Peptidase_C25.
DR Pfam: PF01364; Peptidase_C25; 3.
DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
SQ SEQUENCE 1732 AA: 187914 MW: 45D5B9137791703 CRC64:

Query Match 97.9%: Score 715; DB 2: Length 1732;
Best Local Similarity 97.8%: Pred. No. 2,4e-54;
Matches 131: Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 ADFTETFEESTHGEPAEMTTIDADGEGMCLSSGQDMLTAHGCTNVVSSFSWNGMA 60
DB 1157 ADFTETFEESTHGEPAEMTTIDADGEGMCLSSGQDMLTAHGCTNVVSSFSWNGMA 1216
OY 61 LNPBNYLISKDVYGTATVKYKYYYPVNDGFPGDHYAVMISKTGTNAGDFTVVEETPNCINK 120
DB 1217 LNPBNYLISKDVYGTATVKYKYYAVNDGFPGDHYAVMISKTGTNAGDFTVVEETPNCINK 1276
OY 121 GGARFGLSTEANGA 134
DB 1277 GGARFGLSTEANGA 1290

RESULT 8
ID 051839

```

ID 051839 PRELIMINARY; PRT: 1706 AA.
AC 051839: 051840;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Arginine-specific thiol protease precursor.
GN PRTR.
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; CFB group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
OC Porphyromonas.
NCBI_TaxID=837;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WS0;
RX MEDLINE=95160709; PubMed=7857299;
RA Kirsztbaum L., Solitropoulos C., Jackson C., Cleal S., Slakeski N.,
Reynolds E.C.;
RT "Complete nucleotide sequence of a gene prtr of Porphyromonas
gingivalis W50 encoding a 132 kDa protein that contains an arginine-
specific thiol endopeptidase domain and a haemagglutinin domain.";
RL Biochem. Biophys. Res. Commun. 207:424-431(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=WS0;
RX MEDLINE=96311339; PubMed=8713096;
RA Slakeski N., Cleal S.M., Reynolds E.C.;
RT "Characterization of a Porphyromonas gingivalis gene prtr that encodes
an arginine-specific thiol proteinase and multiple adhesins.";
RL Biochem. Biophys. Res. Commun. 224:605-610(1996).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=WS0;
RX Reynolds E.;
RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=WS0;
RX Slakeski N.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
DR HSSP: P95493; 1CVR.
DR InterPro: IPR001769; Peptidase_C25.
DR Pfam: PF01364; Peptidase_C25.
DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
FT SIGNAL 1 23 POTENTIAL.
FT CHAIN 228 1706 ARGININE-SPECIFIC THIOL PROTEASE.
SQ SEQUENCE 1706 AA: 185626 MW: 4886707C98138844 CRC64;

Query Match 97.8%; Score 714; DB 2; Length 1706;
Best Local Similarity 97.8%; Pred. No. 2,8e-54;
Matches 131; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 ADPTETFESSTHGEAPAEWTTTIDADGGEGWLCSSGQDMLTAHGCTNVVSSFSWNGMA 60
DB 1139 ADPTETFESSTHGEAPAEWTTTIDADGGEGWLCSSGQDMLTAHGCTNVVSSFSWNGMA 1198
OY 61 LNPNDNYLISKDVTGATKVKYKYYVNDGFPDGHVAMISKTGTNAGDFTVVEETPNCINK 120
DB 1199 LNPNDNYLISKDVTGATKVKYKYYVNDGFPDGHVAMISKTGTNAGDFTVVEETPNCINK 1258
OY 121 GGARFGLSTEANGA 134
DB 1259 GGARFGLSTEADGA 1272

```

```

DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Lys-gingipain.
GN KGP.
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; CFB group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
OC Porphyromonas.
NCBI_TaxID=837;
RN [1]
RP SEQUENCE FROM N.A.
RC Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
RX EMBL: U54691; AAA99810.1; -.
DR MEROPS: C25.002; -.
DR InterPro: IPR000977; DNA_Ligase.
DR InterPro: IPR001769; Peptidase_C25.
DR Pfam: PF01364; Peptidase_C25.
DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
SQ SEQUENCE 1723 AA: 186831 MW: 450847E50197CEBD CRC64;

```

```

Query Match 97.4%; Score 711; DB 2; Length 1723;
Best Local Similarity 97.0%; Pred. No. 5.3e-54;
Matches 130; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 ADPTETFESSTHGEAPAEWTTTIDADGGEGWLCSSGQDMLTAHGCTNVVSSFSWNGMA 60
DB 1156 ADPTETFESSTHGEAPAEWTTTIDADGGEGWLCSSGQDMLTAHGCTNVVSSFSWNGMA 1215
OY 61 LNPNDNYLISKDVTGATKVKYKYYVNDGFPDGHVAMISKTGTNAGDFTVVEETPNCINK 120
DB 1216 LNPNDNYLISKDVTGATKVKYKYYVNDGFPDGHVAMISKTGTNAGDFTVVEETPNCINK 1275
OY 121 GGARFGLSTEANGA 134
DB 1276 GGARFGLSTEADGA 1289

```

```

RESULT 10
P96967 PRELIMINARY; PRT: 1358 AA.
ID P96967;
AC P96967;
DT 01-MAY-1997 (TREMblrel. 03, Created)
DT 01-MAY-1997 (TREMblrel. 03, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Hemagglutinin.
GN HAGD.
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; CFB group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
OC Porphyromonas.
NCBI_TaxID=837;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=381;
RA Han N., Lepine G., Whitlock J., Wojciechowski L., Proguishe-Fox A.;
RT "Cloning, sequencing and characterization of hagD, a member of the
HAREP multigene family in Porphyromonas gingivalis.";
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: U68468; AAA49691.1; -.
DR InterPro: IPR000977; DNA_Ligase.
DR InterPro: IPR001769; Peptidase_C25.
DR Pfam: PF01364; Peptidase_C25.
DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
SQ SEQUENCE 1358 AA: 147102 MW: 47FCA0B25B06DE8D CRC64;

```

```

Query Match 97.1%; Score 709; DB 2; Length 1358;
Best Local Similarity 97.0%; Pred. No. 5.9e-54;
Matches 130; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 ADPTETFESSTHGEAPAEWTTTIDADGGEGWLCSSGQDMLTAHGCTNVVSSFSWNGMA 60
DB 791 ADPTETFESSTHGEAPAEWTTTIDADGGEGWLCSSGQDMLTAHGCTNVVSSFSWNGMA 850

```

```

OY 61 LNPBNYLISKDVTGATKVKYKYPVNDGFGPDHYAVMISKGTNAGDFTVVEETPNCINK 120
DB 851 LNPBNYLISKDVTGATKVKYKYPVNDGFGPDHYAVMISKGTNAGDFTVVEETPNCINK 910
OY 121 GGARFGLSTEANGA 134
DB 911 GGARFGLSTEANGA 924

RESULT 11
ID P72194 PRELIMINARY; PRT: 1723 AA.
AC P72194:
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Lys-gingipalain.
GN KGP.
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; CFb group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
OC Porphyromonas.
NCBI_TaxID=837.
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE=97044756; PubMed=8889827;
RC STRAIN=381;
RA Okamoto K., Kadowaki T., Nakayama K., Yamamoto K.;
RT Cloning and sequencing of the gene encoding a novel lysine-specific
RT cysteine proteinase (Lys-gingipalain) in Porphyromonas gingivalis:
RT structural relationship with the arginine-specific cysteine proteinase
RT (Arg-gingipalain).";
RL J. Biochem. 120:398-406(1996).
DR EMBL: D83258; BAA11870.1; -.
DR MEROPS: C25.002; -.
DR InterPro: IPR000977; DNA_Ligase.
DR InterPro: IPR001769; Peptidase_C25.
DR Pfam: PF01364; Peptidase_C25; 3.
DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
SQ SEQUENCE 1723 AA; 187261 MW; 5628963D51493EB CRC64;

Query Match 97.1%; Score 709; DB 2; Length 1723;
Best Local Similarity 97.0%; Pred. No. 7, 9e-54;
Matches 130; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 ADFTETFESSTHGEAPAEWTTIDADGCGWLCSSGQDMLTAHAGTNNVSSFSMNGMA 60
DB 1156 ADFTETFESSTHGEAPAEWTTIDADGCGWLCSSGQDMLTAHAGTNNVSSFSMNGMA 1215
OY 61 LNPBNYLISKDVTGATKVKYKYPVNDGFGPDHYAVMISKGTNAGDFTVVEETPNCINK 120
DB 1216 LNPBNYLISKDVTGATKVKYKYPVNDGFGPDHYAVMISKGTNAGDFTVVEETPNCINK 1275
OY 121 GGARFGLSTEANGA 134
DB 1276 GGARFGLSTEANGA 1289

RESULT 12
ID 051838 PRELIMINARY; PRT: 1706 AA.
AC 051838:
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Protease precursor.
GN PRPR1.
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; CFb group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
OC Porphyromonas.
NCBI_TaxID=837.
RN 11
RP SEQUENCE FROM N.A.

```

```

RC STRAIN-W50:
RA MEDLINE=96071894; PubMed=7591131;
RX Aduse-Opoku J., Muir J., Stanley J.M., Rangarajan M., Curtis M.A.;
RT "Characterization, genetic analysis, and expression of a protease
RT antigen (Prp1) of Porphyromonas gingivalis W50.";
RL Infect. Immun. 63:4744-4754(1995).
RN 12
RP SEQUENCE FROM N.A.
RC STRAIN-W50:
RA Rangarajan M., Aduse-Opoku J., Stanley J.M., Young K.A., Curtis M.A.;
RT "The prp1 and the prp2 arginine-specific protease genes of
RT Porphyromonas gingivalis W50 produce five biochemically distinct
RT enzymes.";
RL Mol. Microbiol. 23:0-0(1997).
DR EMBL: X82680; CA57997.1; -.
DR HSSP: P95493; ICVR.
DR MEROPS: C25.001; -.
DR InterPro: IPR000977; DNA_Ligase.
DR InterPro: IPR001769; Peptidase_C25.
DR Pfam: PF01364; Peptidase_C25; 3.
DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
KW Signal.
FT SIGNAL. 1 23 POTENTIAL.
FT CHAIN 228 719 ALPHA-PROTEASE.
FT CHAIN 720 1262 BETA-ADHESIN.
SQ SEQUENCE 1706 AA; 185705 MW; 0E36DCD87EDA8CDD CRC64;

Query Match 96.7%; Score 706; DB 2; Length 1706;
Best Local Similarity 97.0%; Pred. No. 1, 4e-53;
Matches 130; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 ADFTETFESSTHGEAPAEWTTIDADGCGWLCSSGQDMLTAHAGTNNVSSFSMNGMA 60
DB 1139 ADFTETFESSTHGEAPAEWTTIDADGCGWLCSSGQDMLTAHAGTNNVSSFSMNGMA 1198
OY 61 LNPBNYLISKDVTGATKVKYKYPVNDGFGPDHYAVMISKGTNAGDFTVVEETPNCINK 120
DB 1199 LNPBNYLISKDVTGATKVKYKYPVNDGFGPDHYAVMISKGTNAGDFTVVEETPNCINK 1258
OY 121 GGARFGLSTEANGA 134
DB 1259 GGARFGLSTEANGA 1272

RESULT 13
ID 09FAJ0 PRELIMINARY; PRT: 925 AA.
AC 09FAJ0:
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Putative outer membrane protein PG57.
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; CFb group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
OC Porphyromonas.
NCBI_TaxID=837.
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN-W50:
RA Ross B.C., Barr I., Patterson M., Agius C., Rothel L., Margets M.,
RA Hocking D., Webb E.;
RT "P. gingivalis polypeptides and nucleic acids.";
RT Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
RN 12
RP SEQUENCE FROM N.A.
RC STRAIN-W50:
RA Ross B.C., Czajkowski L., Hocking D., Margets M., Webb E., Rothel L.,
RA Patterson M., Agius C., Camuglia S., Reynolds E., Barr I.G.;
RT "Identification of vaccine candidates from genomic analysis of
RT Porphyromonas gingivalis.";
RT Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY007522; AAG24228.1; -.
DR InterPro: IPR003961; FN_III.

```

DR SMART: SM00060; FN3; 3
 SQ SEQUENCE 925 AA; 103632 MW; 5FF2198D6914DAE8 CRC64;
 Query Match 28.1%; Score 205; DB 2; Length 925;
 Best Local Similarity 37.7%; Pred. No. 1e-09;
 Matches 53; Conservative 14; Mismatches 43; Indels 34; Gaps 7;
 QY 1 ADETFEFSST-----HGEAPAEWTTIDADGDCGMLCLSSGQ---- 38
 Db 288 AAYDESLESTSVCGTLNHYATDAILYENFENGPPVPMGMVVIDADGCGFSW----CHYLNA 342
 QY 39 LDMTLTAGTIVVSSFSW-NGMA-LNPNDNYLSKDYTGATKYKYYVPVND-GFPGDHVAV 95
 Db 343 YDAPPGHNGGCSLSASVPGIGVPTPNYLTTPKVEGAKRVKIVSTODANMAAEHVAV 402
 QY 96 MISKGTNAGDFTVFEET-----PNG 117
 Db 403 MASTGTAVGDFVILFEETMAKPTG 428
 RESULT 14
 ID 09KIB3 PRELIMINARY; PRT; 312 AA.
 AC 09KIB3:
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Hypothetical outer membrane protein Pg27.
 OS Porphyromonas gingivalis (Bacteroides gingivalis).
 OC Bacteria; CFb group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
 OC Porphyromonas.
 CX NCBI_TaxID=837;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=M50;
 RA Ross B., Barr I., Puttersen M., Agius C., Rothel L., Margelits M.,
 RA Hocking D., Webb E.;
 RT "P. gingivalis polypeptides and nucleic acids";
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=M50;
 RA Rosa B. C.;
 RT "Genomic analysis of Porphyromonas gingivalis for vaccine discovery";
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF237555; AAF81413.1; -;
 DR InterPro: IPR001769; Peptidase_C25.
 DR Pfam: PF01364; Peptidase_C25; 2.
 SQ SEQUENCE 312 AA; 34592 MW; 0D5792C9643A25F5 CRC64;
 Query Match 24.7%; Score 180; DB 2; Length 312;
 Best Local Similarity 42.0%; Pred. No. 4.2e-08;
 Matches 50; Conservative 16; Mismatches 35; Indels 18; Gaps 7;
 QY 4 TETFESSTHGEAPAEWTTIDADGDCGMLCLSSGOLDMLTAGTIVVSSFSNMG--MAL 61
 Db 5 SESFESG---IPAIWKITIDADGDGYNMHLTN----FTGQSGL-CVSSASVIGVGAL 54
 QY 62 NPDNYLSKDYVGAT----KVKYYVPVND-GFPGDHVAVMISKGTNAGDF-TVFEET 114
 Db 55 TPDNYLTPELKLPTALVEITVWCTQDLTAPSEHVAVYSSSTGNNAADPVLNLYEET 113
 RESULT 15
 ID 09J215 PRELIMINARY; PRT; 218 AA.
 AC 09J215:
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE Isomerase, putative.
 DE NMB1338.
 OS Neisseria meningitidis (serogroup B).

OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
 OX NCBI_TaxID=491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MC58 / SEROGROUP B;
 RX MEDLINE=20175755; Pubmed=10710307;
 RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
 RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
 RA Nelson W.C., Gwynn M.L., Deboy R., Peterson J.D., Dougherty B.A.,
 RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
 RA Mason T., Ciecko A., Parksey D.S., Blair E., Cifton H., Clark E.B.,
 RA Cotton M.D., Uterback T.R., Khouri H., Qin H., Yatchewyan J.,
 RA Gill J., Scarlato V., Masiagnani V., Pizsa M., Grandi G., Sun L.,
 RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
 RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
 MC58";
 RL Science 287:1809-1815(2000).
 DR EMBL: AE002482; AAF41713.1; -;
 DR TIGR: NMB1338; -;
 DR InterPro: IPR002529; FAA_hydrolase.
 DR Pfam: PF01557; FAA_hydrolase; 1.
 KW Complete proteome.
 SQ SEQUENCE 218 AA; 23058 MW; BA531FCD9C010BD9 CRC64;
 Query Match 12.4%; Score 90.5; DB 16; Length 218;
 Best Local Similarity 26.6%; Pred. No. 2.1;
 Matches 38; Conservative 14; Mismatches 48; Indels 43; Gaps 6;
 QY 23 DADDGEGC-----WCLSSGOLDMLTAGH--GTNVVSSFSNMG 59
 Db 77 DSDGTGECKDLIGCVAGVGLDLTARDIQCLKELKGLPWLKAKGFRHSACVSDFAAGR 136
 QY 60 ALNPDNYLSKDYVGATKVK-----YYVPVNDGFPGDHVAVMISKT----GTNAGDFTV 110
 Db 137 IGNEKVFSLKQNGCVLKRQDGTGLMIVPIRE-----ILHKLADYGLGKGD--LV 185
 QY 111 FEETPNKNGKGFARGLSTEANG 133
 Db 186 FTGTPSGVGAIGAGDNLALELDQ 208

Search completed: June 26, 2003, 00:53:26
 Job time : 41.9882 secs

Thu Jun 20 11:58:19 2003

us-09-980-370-8.ra1

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: June 26, 2003, 00:50:09 ; Search time 3.08824 Seconds

(without alignments)
200.076 Million cell updates/sec

Title: US-09-980-370-8

Perfect score: 108

Sequence: 1 ALMPDNYLSKDYTGATKVKY 21

Scoring table: BLOSUM62

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: Issued_Patents_AA:*
2: /cgn2_6/prodata/1/1aa/5A.COMB.pep:*
3: /cgn2_6/prodata/1/1aa/5B.COMB.pep:*
4: /cgn2_6/prodata/1/1aa/6A.COMB.pep:*
5: /cgn2_6/prodata/1/1aa/6B.COMB.pep:*
6: /cgn2_6/prodata/1/1aa/Backfiles.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	108	100.0	439 2 US-08-570-311-22	Sequence 22, App1
2	108	100.0	450 2 US-08-570-311-16	Sequence 16, App1
3	108	100.0	456 2 US-08-570-311-18	Sequence 18, App1
4	108	100.0	456 2 US-08-570-311-20	Sequence 20, App1
5	108	100.0	1087 2 US-08-570-311-8	Sequence 8, App1
6	108	100.0	1087 2 US-08-570-311-8	Sequence 8, App1
7	108	100.0	1358 2 US-08-570-311-27	Sequence 27, App1
8	108	100.0	1687 2 US-08-570-311-29	Sequence 29, App1
9	108	100.0	1704 3 US-08-336-308A-10	Sequence 10, App1
10	108	100.0	1704 4 US-09-490-931-10	Sequence 10, App1
11	108	100.0	1732 2 US-08-570-311-10	Sequence 10, App1
12	108	100.0	1732 2 US-08-353-485-10	Sequence 10, App1
13	108	100.0	2628 2 US-08-570-311-14	Sequence 14, App1
14	108	100.0	2628 2 US-08-570-311-14	Sequence 14, App1
15	47	43.5	183 4 US-09-117-257-42	Sequence 42, App1
16	47	43.5	183 4 US-09-489-352-42	Sequence 42, App1
17	47	43.5	187 2 US-08-589-711-2	Sequence 2, App1
18	47	43.5	187 4 US-09-117-257-8	Sequence 8, App1
19	47	43.5	187 4 US-09-117-257-9	Sequence 9, App1
20	47	43.5	187 4 US-09-117-257-15	Sequence 15, App1
21	47	43.5	187 4 US-09-117-257-23	Sequence 23, App1
22	47	43.5	187 4 US-09-117-257-25	Sequence 25, App1
23	47	43.5	187 4 US-09-117-257-32	Sequence 32, App1
24	47	43.5	187 4 US-09-221-938-2	Sequence 2, App1
25	47	43.5	187 4 US-08-945-476-9	Sequence 9, App1
26	47	43.5	187 4 US-08-945-476-15	Sequence 15, App1
27	47	43.5	187 4 US-08-945-476-23	Sequence 23, App1

28	47	43.5	187 4 US-08-945-476-25	Sequence 25, App1
29	47	43.5	187 4 US-08-489-352-8	Sequence 8, App1
30	47	43.5	187 4 US-08-489-352-9	Sequence 9, App1
31	47	43.5	187 4 US-08-489-352-15	Sequence 15, App1
32	47	43.5	187 4 US-08-489-352-23	Sequence 23, App1
33	47	43.5	187 4 US-08-489-352-25	Sequence 25, App1
34	47	43.5	187 4 US-08-489-352-32	Sequence 32, App1
35	47	43.5	188 4 US-09-117-257-27	Sequence 27, App1
36	47	43.5	188 4 US-09-117-257-30	Sequence 30, App1
37	47	43.5	188 4 US-08-945-476-27	Sequence 27, App1
38	47	43.5	188 4 US-08-489-352-27	Sequence 27, App1
39	47	43.5	188 4 US-08-489-352-30	Sequence 30, App1
40	47	43.5	188 4 US-09-117-257-44	Sequence 44, App1
41	47	43.5	188 4 US-09-117-257-36	Sequence 36, App1
42	47	43.5	191 4 US-08-489-352-36	Sequence 36, App1
43	47	43.5	191 4 US-08-570-311-2	Sequence 2, App1
44	46	42.6	497 2 US-08-353-485-2	Sequence 2, App1
45	46	42.6	497 2 US-08-353-485-2	Sequence 2, App1

ALIGNMENTS

RESULT 1
US-08-570-311-22
Sequence 22, Application US/08570311
Patent No. 5824791
GENERAL INFORMATION:
APPLICANT: Prologis-Fox, Ann
APPLICANT: Tumwasorn, Soumying
APPLICANT: Leptine, GuyYalene
APPLICANT: Han, Naiming
APPLICANT: Lantzi, Marilyn
APPLICANT: Patti, Joseph
TITLE OF INVENTION: Cloned Porphyromonas gingivalis genes
TITLE OR INVENTION: and Probes for the Detection of Periodontal Disease
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ted W. Whitlock
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/570.311
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/353,485
FILING DATE: 09-DEC-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/647,119
FILING DATE: 25-JAN-1991
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/241,640
FILING DATE: 08-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UP15.C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:

Thu Jun 26 11:58:19 2003

us-09-980-370-8.ra1

Page 2

LENGTH: 439 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-570-311-22

Query Match 100.0%; Score 108; DB 2; Length 439.
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNPDNYLSKDVGTGATKVKY 21
DB 251 ALNPDNYLSKDVGTGATKVKY 271

RESULT 2
US-08-570-311-16

; Sequence 16, Application US/08570311
; Patent No. 5824791
; GENERAL INFORMATION:
; APPLICANT: Proguiske-Fox, Ann
; APPLICANT: Tumwasorn, Somying
; APPLICANT: Lepine, Guylaine
; APPLICANT: Han, Naiming
; APPLICANT: Lantzi, Marilyn
; APPLICANT: Pate, Joseph
; TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
; TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ted W. Whitlock
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/570,311
; FILING DATE: 25-JAN-1991
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/241,640
; FILING DATE: 09-DEC-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/647,119
; FILING DATE: 25-JAN-1991
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/241,640
; FILING DATE: 08-SEP-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Whitlock, Ted W.
; REGISTRATION NUMBER: 36,965
; REFERENCE/DOCKET NUMBER: UP15.C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (904) 372-5800
; TELEFAX: (904) 372-5800
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 450 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-570-311-16

Query Match 100.0%; Score 108; DB 2; Length 450;
Best Local Similarity 100.0%; Pred. No. 2.1e-10;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNPDNYLSKDVGTGATKVKY 21
DB 245 ALNPDNYLSKDVGTGATKVKY 265

RESULT 3
US-08-570-311-18

; Sequence 18, Application US/08570311
; Patent No. 5824791
; GENERAL INFORMATION:
; APPLICANT: Proguiske-Fox, Ann
; APPLICANT: Tumwasorn, Somying
; APPLICANT: Lepine, Guylaine
; APPLICANT: Han, Naiming
; APPLICANT: Lantzi, Marilyn
; APPLICANT: Pate, Joseph
; TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
; TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ted W. Whitlock
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/570,311
; FILING DATE: 25-JAN-1991
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/353,485
; FILING DATE: 09-DEC-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/647,119
; FILING DATE: 25-JAN-1991
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/241,640
; FILING DATE: 08-SEP-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Whitlock, Ted W.
; REGISTRATION NUMBER: 36,965
; REFERENCE/DOCKET NUMBER: UP15.C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (904) 372-5800
; TELEFAX: (904) 372-5800
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 456 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-570-311-18

Query Match 100.0%; Score 108; DB 2; Length 456;
Best Local Similarity 100.0%; Pred. No. 2.1e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNPDNYLSKDVGTGATKVKY 21
DB 251 ALNPDNYLSKDVGTGATKVKY 271

RESULT 4

US-08-570-311-20

Sequence 20, Application US/08570311

Patent No. 5824791

GENERAL INFORMATION:

APPLICANT: Proguiske-Fox, Ann

APPLICANT: Tumwasorn, Somying

APPLICANT: Lepline, Guyalaine

APPLICANT: Han, Naïming

APPLICANT: Lantze, Marilyn

APPLICANT: Patei, Joseph

TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes

TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease

NUMBER OF SEQUENCES: 29

CORRESPONDENCE ADDRESS:

ADDRESSER: Ted W. Whitlock

STREET: 2421 N.W. 41st Street, Suite A-1

CITY: Gainesville

STATE: FL

COUNTRY: USA

ZIP: 32606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/570,311

FILING DATE: 08-SEP-1988

CLASSIFICATION: 424

PRIORITY APPLICATION DATA:

APPLICATION NUMBER: US 07/241,640

FILING DATE: 08-SEP-1988

ATTORNEY/AGENT INFORMATION:

NAME: Whitlock, Ted W.

REGISTRATION NUMBER: 36,965

REFERENCE/DOCKET NUMBER: UF15 C3

TELECOMMUNICATION INFORMATION:

TELEPHONE: (904) 375-8100

TELEFAX: (904) 372-5800

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:

LENGTH: 456 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-570-311-20

Query Match

Best Local Similarity 100.0%; Score 108; DB 2; Length 456;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALNPDNYLISKDYVGATKRY 21

DB 251 ALNPDNYLISKDYVGATKRY 271

RESULT 5

US-08-570-311-8

Sequence 8, Application US/08570311

Patent No. 5824791

GENERAL INFORMATION:

APPLICANT: Proguiske-Fox, Ann

APPLICANT: Tumwasorn, Somying

APPLICANT: Lepline, Guyalaine

APPLICANT: Han, Naïming

APPLICANT: Lantze, Marilyn

APPLICANT: Patei, Joseph

TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes

TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease

NUMBER OF SEQUENCES: 29

CORRESPONDENCE ADDRESS:

ADDRESSER: Ted W. Whitlock

STREET: 2421 N.W. 41st Street, Suite A-1

CITY: Gainesville

STATE: FL

COUNTRY: USA

ZIP: 32606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/570,311

FILING DATE: 08-SEP-1988

CLASSIFICATION: 424

PRIORITY APPLICATION DATA:

APPLICATION NUMBER: US 07/241,640

FILING DATE: 08-SEP-1988

ATTORNEY/AGENT INFORMATION:

NAME: Whitlock, Ted W.

REGISTRATION NUMBER: 36,965

REFERENCE/DOCKET NUMBER: UF15 C3

TELECOMMUNICATION INFORMATION:

TELEPHONE: (904) 375-8100

TELEFAX: (904) 372-5800

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 1087 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-570-311-8

Query Match

Best Local Similarity 100.0%; Score 108; DB 2; Length 1087;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALNPDNYLISKDYVGATKRY 21

DB 579 ALNPDNYLISKDYVGATKRY 599

RESULT 6

US-08-353-485-8

Sequence 8, Application US/08353485

Patent No. 5830710

GENERAL INFORMATION:

APPLICANT: Proguiske-Fox, Ann

APPLICANT: Tumwasorn, Somying

APPLICANT: Lepline, Guyalaine

APPLICANT: Han, Naïming

APPLICANT: Lantze, Marilyn

APPLICANT: Patei, Joseph

TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes

TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSER: Ted W. Whitlock

STREET: 2421 N.W. 41st Street, Suite A-1

CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/353,485
FILING DATE: 09-DEC-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/647,119
FILING DATE: 25-JAN-1991
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/241,640
FILING DATE: 08-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UF15.C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1087 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-353-485-8

Query Match 100.0%; Score 108; DB 2; Length 1087;
Best Local Similarity 100.0%; Pred. No. 6.3e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNPDNVLISKDYTGATKVKY 21
|||||
DB 579 ALNPDNVLISKDYTGATKVKY 599

RESULT 7
US-08-570-311-27
Sequence 27, Application US/08570311
Patent No. 5824791
GENERAL INFORMATION:
APPLICANT: Proguiske-Fox, Ann
APPLICANT: Tumwasorn, Somying
APPLICANT: Lepine, Guyline
APPLICANT: Han, Naiming
APPLICANT: Lantiz, Marilyn
APPLICANT: Patti, Joseph
TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ted W. Whitlock
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/570,311
FILING DATE: 08/570,311

CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/353,485
FILING DATE: 09-DEC-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/647,119
FILING DATE: 25-JAN-1991
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/241,640
FILING DATE: 08-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UF15.C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 1358 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-570-311-27

Query Match 100.0%; Score 108; DB 2; Length 1358;
Best Local Similarity 100.0%; Pred. No. 8.3e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNPDNVLISKDYTGATKVKY 21
|||||
DB 850 ALNPDNVLISKDYTGATKVKY 870

RESULT 8
US-08-570-311-29
Sequence 29, Application US/08570311
Patent No. 5824791
GENERAL INFORMATION:
APPLICANT: Proguiske-Fox, Ann
APPLICANT: Tumwasorn, Somying
APPLICANT: Lepine, Guyline
APPLICANT: Han, Naiming
APPLICANT: Lantiz, Marilyn
APPLICANT: Patti, Joseph
TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ted W. Whitlock
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/570,311
FILING DATE: 08/570,311
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/353,485
FILING DATE: 09-DEC-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/647,119
FILING DATE: 25-JAN-1991

CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/241,640
FILING DATE: 08-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UP15.C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 1687 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-570-311-29

Query Match 100.0%; Score 108; DB 2; Length 1687;
Best Local Similarity 100.0%; Pred. No. 1,1e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNPDNYLISKDVGTGATKVKY 21
DB 1179 ALNPDNYLISKDVGTGATKVKY 1199

RESULT 9
US-08-336-308A-10
Sequence 10, Application US/08336308A
Patent No. 6017532
GENERAL INFORMATION:
APPLICANT: Travis, James
APPLICANT: Potempa, Jan S.
APPLICANT: Barry, Philip J.
APPLICANT: Pavloff, Nadine
TITLE OF INVENTION: Porphyromonas gingivalis
TITLE OF INVENTION: Arginine-specific Proteinase Coding Sequences
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: US
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/336,308A
FILING DATE: 08-NOV-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/119,361
FILING DATE: 10-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/265,441
FILING DATE: 24-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Feibert, Donna M.
REGISTRATION NUMBER: 33,878
REFERENCE/DOCKET NUMBER: 21-93C
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 499-8080
TELEFAX: (303) 499-8089
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1704 amino acids
TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-336-308A-10

Query Match 100.0%; Score 108; DB 3; Length 1704;
Best Local Similarity 100.0%; Pred. No. 1,1e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNPDNYLISKDVGTGATKVKY 21
DB 1196 ALNPDNYLISKDVGTGATKVKY 1216

RESULT 10
US-08-822-324-6
Sequence 6, Application US/08822324
Patent No. 6129917
GENERAL INFORMATION:
APPLICANT: Potempa, Jan S.
APPLICANT: Travis, James
APPLICANT: Genco, Caroline A.
TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS COMPRISING
TITLE OF INVENTION: PORPHYROMONAS GINGIVALIS PROTEINS AND/OR PEPTIDES AND
TITLE OF INVENTION: METHODS
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: CO
COUNTRY: US
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/822,324
FILING DATE: 21-MAR-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/013,945
FILING DATE: 22-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Feibert, Donna M.
REGISTRATION NUMBER: 33,878
REFERENCE/DOCKET NUMBER: 103-95 WO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 488-8080
TELEFAX: (303) 499-8089
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1704 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-822-324-6

Query Match 100.0%; Score 108; DB 3; Length 1704;
Best Local Similarity 100.0%; Pred. No. 1,1e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNPDNYLISKDVGTGATKVKY 21
DB 1196 ALNPDNYLISKDVGTGATKVKY 1216

RESULT 11
US-09-480-931-10
Sequence 10, Application US/09490931
Patent No. 6274718
GENERAL INFORMATION:

APPLICANT: Travis, James
APPLICANT: Potempa, Jan S.
APPLICANT: Bart, Philip J.
APPLICANT: Pavloff, Nadine
TITLE OF INVENTION: Porphyromonas gingivalis
TITLE OF INVENTION: Arginine-specific Proteinase Coding Sequences
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: US
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/490,931
CLASSIFICATION:
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/336,308
FILING DATE:
PRIOR APPLICATION NUMBER:
APPLICATION NUMBER: US 08/265,441
FILING DATE: 24-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Felder, Donna M.
REGISTRATION NUMBER: 33,878
REFERENCE/DOCKET NUMBER: 21-93C
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 499-8080
TELEFAX: (303) 499-8089
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1704 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-490-931-10
Query Match
Best Local Similarity 100.0%; Score 108; DB 4; Length 1704;
Matches 21: Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ALNPDNYLSKDVGTGATKVKY 21
DB 1196 ALNPDNYLSKDVGTGATKVKY 1216
RESULT 12
US-08-570-311-10
Sequence 10, Application US/08570311
Patent No. 5824791
GENERAL INFORMATION:
APPLICANT: Proguiske-Fox, Ann
APPLICANT: Tumwasorn, Somying
APPLICANT: Lepling, Guylaiane
APPLICANT: Han, Naiming
APPLICANT: Lantz, Marilynn
APPLICANT: Patti, Joseph
TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ted W. Whitlock
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA

ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/570,311
CLASSIFICATION: 424
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/353,485
FILING DATE: 09-DEC-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/647,119
FILING DATE: 25-JAN-1991
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/241,640
FILING DATE: 08-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UF15.C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1732 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-570-311-10
Query Match
Best Local Similarity 100.0%; Score 108; DB 2; Length 1732;
Matches 21: Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ALNPDNYLSKDVGTGATKVKY 21
DB 1216 ALNPDNYLSKDVGTGATKVKY 1236
RESULT 13
US-08-353-485-10
Sequence 10, Application US/08353485
Patent No. 5830710
GENERAL INFORMATION:
APPLICANT: Proguiske-Fox, Ann
APPLICANT: Tumwasorn, Somying
APPLICANT: Lepling, Guylaiane
APPLICANT: Han, Naiming
APPLICANT: Lantz, Marilynn
APPLICANT: Patti, Joseph
TITLE OF INVENTION: Cloned Porphyromonas gingivalis Genes
TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ted W. Whitlock
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/353,485

FILING DATE: 09-DEC-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/647,119
FILING DATE: 25-JAN-1991
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/241,640
FILING DATE: 08-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UFI5.C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1732 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-353-485-10

Query Match 100.0%; Score 108; DB 2; Length 1732;
Best Local Similarity 100.0%; Pred. No. 1.1e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALNPDNYLSKDVGTGATKVKY 21
DB 1216 ALNPDNYLSKDVGTGATKVKY 1236

RESULT 14
US-08-570-311-14
Sequence 14, Application US/08570311
Patent No. 5824791
GENERAL INFORMATION:
APPLICANT: Proguis-Kee-Fox, Ann
APPLICANT: Tumwasorn, Somying
APPLICANT: Leplue, Guyaline
APPLICANT: Han, Nalinng
APPLICANT: Lantz, Marilyn
APPLICANT: Patti, Joseph
TITLE OF INVENTION: Cloned Porphyrinomas gingivalls Genes
TITLE OF INVENTION: and Probes for the Detection of Periodontal Disease
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ted W. Whitlock
STREET: 2421 N.W. Whitlock
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/570,311
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/353,485
FILING DATE: 09-DEC-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/647,119
FILING DATE: 25-JAN-1991
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/241,640

FILING DATE: 08-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UFI5.C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 2628 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-570-311-14

Query Match 100.0%; Score 108; DB 2; Length 2628;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALNPDNYLSKDVGTGATKVKY 21
DB 744 ALNPDNYLSKDVGTGATKVKY 764

RESULT 15
US-09-117-257-42
Sequence 42, Application US/09117257
Patent No. 6214355
GENERAL INFORMATION:
APPLICANT: Hook, Magnus
APPLICANT: Guo, Betty
APPLICANT: Hanson, Mark
TITLE OF INVENTION: DDBA AND DDBB COMPOSITIONS AND METHODS OF USE
FILE REFERENCE: 4210.000500
CURRENT APPLICATION NUMBER: US/09/117,257
EARLIER FILING DATE: 1998-07-22
EARLIER APPLICATION NUMBER: PCT/US96/17081
EARLIER FILING DATE: 1996-10-22
EARLIER APPLICATION NUMBER: 08/589,711
EARLIER FILING DATE: 1996-01-22
EARLIER APPLICATION NUMBER: 08/427,023
EARLIER FILING DATE: 1995-04-24
NUMBER OF SEQ ID NOS: 66
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO: 42
LENGTH: 183
TYPE: PRT
ORGANISM: Borrelia burgdorferi
US-09-117-257-42

Query Match 43.5%; Score 47; DB 4; Length 183;
Best Local Similarity 66.7%; Pred. No. 1.9;
Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 6 NYLSKDVGTGATKVK 20
DB 20 NLLISGIGTGATKVK 34

Search completed: June 26, 2003, 00:53:58
Job time : 3.08824 secs

Thu Jun 26 11:58:19 2003

us-09-980-370-8.rapb

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 26, 2003, 00:50:39 ; Search time 4.44706 Seconds

(without alignments)
510.977 Million cell updates/sec

Title: US-09-980-370-8

Perfect score: 108

Sequence: 1 ALNPNTYLSKDYTGATKVKY 21

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 417779 seqs, 108206813 residues

Total number of hits satisfying chosen parameters: 417779

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Published Applications-AA:*

- 1: /cgn2_6/prodata/1/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/prodata/1/pubpaa/PCIT_NEW_PUB.pep.*
- 3: /cgn2_6/prodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/prodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/prodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/prodata/1/pubpaa/PCITUS_PUBCOMB.pep.*
- 8: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/prodata/1/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/prodata/1/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/prodata/1/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/prodata/1/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/prodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50	46.3	580	9 US-09-928-457-37	Sequence 37, App1
2	44	40.7	81	10 US-09-529-063-19	Sequence 19, App1
3	44	40.7	575	9 US-09-738-626-4263	Sequence 4263, App1
4	43	39.8	381	10 US-09-815-242-10962	Sequence 10962, App1
5	43	39.8	480	9 US-09-464-099A-49	Sequence 49, App1
6	43	39.8	480	10 US-09-861-696-49	Sequence 49, App1
7	43	39.8	1588	9 US-10-214-766-34	Sequence 34, App1
8	42.5	39.4	369	9 US-09-804-291-457	Sequence 457, App1
9	42.5	39.4	369	10 US-09-886-055-457	Sequence 457, App1
10	42	38.9	2548	10 US-09-851-682A-1	Sequence 1, App1
11	41.5	38.4	491	10 US-09-925-299-1177	Sequence 1177, App1
12	41	38.0	166	9 US-09-881-752A-128	Sequence 128, App1
13	41	38.0	166	10 US-09-925-299-1177	Sequence 1177, App1
14	41	38.0	431	9 US-09-860-670-117	Sequence 117, App1
15	41	38.0	773	9 US-09-808-602-60	Sequence 60, App1
16	41	38.0	773	9 US-09-800-198-49	Sequence 49, App1
17	41	38.0	850	9 US-09-808-602-58	Sequence 58, App1
18	41	38.0	850	9 US-09-800-198-47	Sequence 47, App1
19	41	38.0	983	9 US-09-808-602-73	Sequence 73, App1

20	41	38.0	983	9 US-10-013-136-2	Sequence 2, App1
21	41	38.0	1572	9 US-10-037-182-20	Sequence 20, App1
22	41	38.0	1576	9 US-10-037-182-16	Sequence 16, App1
23	41	38.0	1605	9 US-10-037-182-18	Sequence 18, App1
24	41	38.0	1607	10 US-09-938-275-10	Sequence 10, App1
25	41	38.0	1609	9 US-10-037-182-14	Sequence 14, App1
26	41	38.0	1609	9 US-10-037-182-12	Sequence 12, App1
27	41	38.0	1609	10 US-09-938-275-11	Sequence 11, App1
28	40.5	37.5	574	10 US-09-923-9980-27	Sequence 27, App1
29	40.5	37.5	524	10 US-09-923-9980-21	Sequence 21, App1
30	40.5	37.5	524	10 US-09-923-9980-49	Sequence 49, App1
31	40.5	37.5	524	10 US-09-923-9980-12	Sequence 12, App1
32	40	37.0	1765	9 US-10-037-182-8	Sequence 8, App1
33	40	37.0	1765	9 US-10-037-182-6	Sequence 6, App1
34	40	37.0	1786	9 US-10-037-182-10	Sequence 10, App1
35	40	37.0	1786	9 US-09-938-275-6	Sequence 6, App1
36	40	37.0	1786	10 US-09-938-275-7	Sequence 7, App1
37	40	37.0	1786	10 US-10-037-182-11	Sequence 11, App1
38	40	37.0	1786	10 US-10-037-182-13	Sequence 13, App1
39	39.5	36.6	158	9 US-10-106-698-6717	Sequence 6717, App1
40	39.5	36.6	324	10 US-09-815-242-11345	Sequence 11345, App1
41	39.5	36.6	349	9 US-10-270-875-57	Sequence 57, App1
42	39.5	36.6	349	9 US-10-270-878-57	Sequence 57, App1
43	39.5	36.6	349	9 US-10-270-786-57	Sequence 57, App1
44	39.5	36.6	349	9 US-10-270-710-57	Sequence 57, App1
45	39.5	36.6	349	9 US-10-270-859-57	Sequence 57, App1

ALIGNMENTS

RESULT 1
US-09-928-457-37
Sequence 37, Application US/09928457
Patent No. US20020156403A1
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: DNA, specific proteins and peptides
TITLE OF INVENTION: of the Neisseria meningitidis species bacteria, method
NUMBER OF SEQUENCES: 99
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patentin Release #1.0, Version #1.30 (OEB)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/928.457
FILING DATE: 2001-08-14
PRIORITY APPLICATION NUMBER: 09/214,759
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 580 acids amin.s
TYPE: acids amin.s
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Protein
LOCATION: 1..580
US-09-928-457-37
Query Match: 46.3% Score 50: DB 9: Length 580:
Best local similarity 45.0% Pred. No. 5.3:
Matches 9: Conservative 5: Mismatches 6: Indels 0: Gaps 0:
Gy 2 LNPNNTYLSKDYTGATKVKY 21
Db 56 LEKNNTYLSKDYTGATKVKY 75

RESULT 2
 US-09-529-063-19
 Sequence 19, Application US/09529063
 Patent No. US20020102542A1
 GENERAL INFORMATION:
 APPLICANT: FUKUSHIMA, DAICHI
 APPLICANT: SHIBAYAMA, SHIRO
 APPLICANT: TADA, HIDEAKI
 TITLE OF INVENTION: POLYPEPTIDE, CDNA ENCODING THE POLYPEPTIDE, AND USE OF
 TITLE OF INVENTION: THE BOTH
 FILE REFERENCE: 058769
 CURRENT APPLICATION NUMBER: US/09/529,063
 CURRENT FILING DATE: 2000-04-07
 PRIOR APPLICATION NUMBER: PCT/JP98/04514
 PRIOR FILING DATE: 1998-10-06
 PRIOR APPLICATION NUMBER: JP 9-274674
 PRIOR FILING DATE: 1997-10-07
 NUMBER OF SEQ ID NOS: 117
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 19
 LENGTH: 81
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-529-063-19

```

Query Match      40.78; Score 44; DB 10; Length 81;
Best Local Similarity 53.33; Pred. No. 5;
Matches      8; Conservative 3; Mismatches 4; Indels 0; Gaps 0
OY      4 PDNYLISKDYTGATK 18
      |:|:| |:| |:|
Db      28 PNYSLSSKKIAGLTK 42

```

```

RESULT 3
US-09-738-626-4263
: Sequence 4263, Application US/09738626
: Publication No. US20020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, MAOKO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO
: TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
: FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738, 626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 4263
:
: LENGTH: 575
:
: TYPE: PRT
: ORGANISM: Corynebacterium glutamicum
US-09-738-626-4263

```

Query Match	40.7%	Score 44	DB 9	Length 575
Best Local Similarity	42.96	Pred. NO. 52		
Matches	9	Conservative	4	Mismatches 8
				Indels 0
				Gaps 0
QY	1	ALPNPNVILSKDYGATKVKY	21	
		: : :		

Db 91 ALNPVPHLVPEDETGIGRALY 111

```

US-09-815-242-10962
/US-09-815-242-10962
/Sequence 10962, Application US/09815242
/ Patent No. US20020061569A1
/GENERAL INFORMATION:
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlson, Karl L.
/ APPLICANT: Zyskind, Judith W.
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John D.
/ APPLICANT: Carr, Grant J.
/ APPLICANT: Yamamoto, Robert T.
/ APPLICANT: Xu, H. Howard
/ TITLE OF INVENTION: Identification of Essential Genes In
/ FILE REFERENCE: ELIURA 011A
/ FILE REFERENCE: Prokaryotes
/ CURRENT FILING DATE: 2001-03-21
/ PRIOR APPLICATION NUMBER: US/09/815,242
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-21
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ NUMBER OF SEQ ID NOS: 14110
/ SOFTWARE: PASTE0 for Windows Version 4.0
/ SEQ ID NO 10962
/ LENGTH: 381
/ TYPE: PRT
/ ORGANISM: Haemophilus Influenzae
/US-09-815-242-10962

```

Query Match	39.88;	Score 43;	DB 10;	Length 381;
Best Local Similarity	50.08;	Pred. No. 47;		
Matches	7;	Conservative	3;	Mismatches 4;
				Indels 0;
				Gaps 0;

OY 4 PONTLSKDYTCAT 17
| : | | |
Db 104 PDIEWISSDYGCT 117

RESULT 5
US-99-464-099A--49
Sequence 49, Application US/094464099A
Patent No. US20020168680A1
GENERAL INFORMATION:
APPLICANT: Barry, Gerard F.
APPLICANT: Kishore, Ganesh M.
APPLICANT: Padgett, Stephen R.
APPLICANT: Stallings, William C.
TITLE OF INVENTION: GLYPHOSATE TOLERANT 5-ENOLPYRUVYLSHIKIMATE-3-PHOSPHATE SYNTHAS
FILE REFERENCE: 11899.0175.CNVS01 MOBT:1175-2
CURRENT APPLICATION NUMBER: US/09/464,099A
CURRENT FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: US 09/137,440
PRIOR FILING DATE: 1998-08-20
PRIOR APPLICATION NUMBER: US 08/833,485
PRIOR FILING DATE: 1997-04-07
PRIOR APPLICATION NUMBER: US 08/306,063
PRIOR FILING DATE: 1994-09-13
PRIOR APPLICATION NUMBER: US 07/749,611
PRIOR FILING DATE: 1991-08-28

PRIOR APPLICATION NUMBER: US 07/576,537
PRIOR FILING DATE: 1990-08-31
NUMBER OF SEQ ID NOS: 70
SOFTWARE: PatentIn version 3.0
SEQ ID NO 49
LENGTH: 480
TYPE: PRT
ORGANISM: Saccharomyces cerevisiae
US-09-464-099A-49

Query Match 39.8%; Score 43; DB 9; Length 480;
Best Local Similarity 43.8%; Pred. No. 62;
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 LNPDMYLSKDVTCAT 17
: 11 1:1 1:1
Db 247 INPSEYVIESDASSAT 262

RESULT 6
US-09-861-696-49
Sequence 49, Application US/09861696
Patent No. US2002007053A1
GENERAL INFORMATION:
APPLICANT: Barry, Gerard F.
APPLICANT: Padgett, Stephen R.
APPLICANT: Stallings, William C.
TITLE OF INVENTION: GLYPHOSATE TOLERANT 5-ENOLPYRUVYLSHIKIMATE-3-PHOSPHATE SYNTHASES
FILE REFERENCE: 11899.0175.CNUS04 MOB:175-4
CURRENT APPLICATION NUMBER: US/09/861,696
CURRENT FILING DATE: 2001-05-21
PRIOR APPLICATION NUMBER: US 09/137,440
PRIOR FILING DATE: 1998-08-20
PRIOR APPLICATION NUMBER: US 08/833,485
PRIOR FILING DATE: 1997-04-07
PRIOR APPLICATION NUMBER: US 08/306,063
PRIOR FILING DATE: 1994-09-13
PRIOR APPLICATION NUMBER: US 07/749,611
PRIOR FILING DATE: 1991-08-28
PRIOR APPLICATION NUMBER: US 07/576,537
PRIOR FILING DATE: 1990-08-31
NUMBER OF SEQ ID NOS: 70
SOFTWARE: PatentIn version 3.0
SEQ ID NO 49
LENGTH: 480
TYPE: PRT
ORGANISM: Saccharomyces cerevisiae
US-09-861-696-49

Query Match 39.8%; Score 43; DB 10; Length 480;
Best Local Similarity 43.8%; Pred. No. 62;
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 LNPDMYLSKDVTCAT 17
: 11 1:1 1:1
Db 247 INPSEYVIESDASSAT 262

RESULT 7
US-10-214-766-34
Sequence 34, Application US/10214766
Publication No. US20030084473A1
GENERAL INFORMATION:
APPLICANT: Gocal, Greg
TITLE OF INVENTION: NON-TRANSGENIC HERBICIDE RESISTANT PLANTS
FILE REFERENCE: CA1138
CURRENT APPLICATION NUMBER: US/10/214,766
CURRENT FILING DATE: 2002-08-09
PRIOR APPLICATION NUMBER: US 60/311,734
PRIOR FILING DATE: 2001-08-09
NUMBER OF SEQ ID NOS: 53
SOFTWARE: PatentIn version 3.2

SEQ ID NO 34
LENGTH: 1588
TYPE: PRT
ORGANISM: Saccharomyces cerevisiae
US-10-214-766-34

Query Match 39.8%; Score 43; DB 9; Length 1588;
Best Local Similarity 43.8%; Pred. No. 2,66+02;
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 LNPDMYLSKDVTCAT 17
: 11 1:1 1:1
Db 645 INPSEYVIESDASSAT 660

RESULT 8
US-09-804-291-457
Sequence 457, Application US/09804291
Publication No. US20030088059A1
GENERAL INFORMATION:
APPLICANT: ZOZULIA, SERGEY
TITLE OF INVENTION: HUMAN OLFACTORY RECEPTORS AND GENES ENCODING SAME
FILE REFERENCE: P 0278005
CURRENT APPLICATION NUMBER: US/09/804,291
CURRENT FILING DATE: 2001-03-13
PRIOR APPLICATION NUMBER: 60/188,914
PRIOR FILING DATE: 2000-03-13
PRIOR APPLICATION NUMBER: 60/192,033
PRIOR FILING DATE: 2000-03-24
PRIOR APPLICATION NUMBER: 60/198,474
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/199,335
PRIOR FILING DATE: 2000-04-24
PRIOR APPLICATION NUMBER: 60/207,702
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/213,849
PRIOR FILING DATE: 2000-06-23
PRIOR APPLICATION NUMBER: 60/226,534
PRIOR FILING DATE: 2000-08-16
PRIOR APPLICATION NUMBER: 60/230,732
PRIOR FILING DATE: 2000-09-07
PRIOR APPLICATION NUMBER: 60/266,862
PRIOR FILING DATE: 2001-02-07
NUMBER OF SEQ ID NOS: 529
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 457
LENGTH: 369
TYPE: PRT
ORGANISM: Homo sapiens
US-09-804-291-457

Query Match 39.4%; Score 42.5; DB 9; Length 369;
Best Local Similarity 66.7%; Pred. No. 55;
Matches 12; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

QY 2 LNPDMYLSKDVTCAT 18
: 11 1:1 1:1
Db 335 LNPYLSLRNKDVTCAT 352

RESULT 9
US-09-886-055-457
Sequence 457, Application US/09886055
Patent No. US20020132273A1
GENERAL INFORMATION:
APPLICANT: STRYER, LUBERT
APPLICANT: ZOZULIA, SERGEY
TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND
FILE REFERENCE: 078003-0277150
CURRENT APPLICATION NUMBER: US/09/886,055
CURRENT FILING DATE: 2001-06-22
PRIOR APPLICATION NUMBER: 60/213,812

```

: PRIOR FILING DATE: 2000-06-22
: NUMBER OF SEQ ID NOS: 522
: SOFTWARE: Patent In Ver. 2.1
: SEQ ID NO 457
: LENGTH: 369
: TYPE: PRF
: ORGANISM: Homo sapiens
US-09-886-055-457

Query Match
Best Local Similarity 39.4%; Score 42.5; DB 10; Length 369;
Matches 12; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

QY 2 LNPDMY-LISKDYTGATK 18
DB 335 LNPDMYLRNKKDYTGALK 352

RESULT 10
US-09-851-682A-1
: Sequence 1, Application US/09851682A
: Patent No. US20020091248A1
: GENERAL INFORMATION:
: APPLICANT: Adams, Arwen E.
: APPLICANT: Chiu, Choi Ying
: APPLICANT: Duhl, David
: APPLICANT: Gorman, Susan W.
: APPLICANT: Leng, Song
: APPLICANT: Sheffield, Val
: TITLE OF INVENTION: MYOSIN IXA AND CYCLIC NUCLEOTIDE GATED
: TITLE OF INVENTION: CHANNEL-15 (CNGC-15) POLYNUCLEOTIDES, POLYPEPTIDES,
: TITLE OF INVENTION: COMPOSITIONS, METHODS, AND USES THEREOF
: FILE REFERENCE: 200130.442
: CURRENT APPLICATION NUMBER: US/09/851,682A
: CURRENT FILING DATE: 2001-05-08
: PRIOR APPLICATION NUMBER: US/09/172,422
: PRIOR FILING DATE: 1998-10-14
: NUMBER OF SEQ ID NOS: 3
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 1
: LENGTH: 2548
: TYPE: PRF
: ORGANISM: Homo sapien
US-09-851-682A-1

Query Match
Best Local Similarity 38.9%; Score 42; DB 10; Length 2548;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 LNPDMY-LISK 11
DB 989 LNPDMYQVGR 998

RESULT 11
US-09-881-752A-128
: Sequence 128, Application US/09881752A
: Patent No. US20020115078A1
: GENERAL INFORMATION:
: APPLICANT: Kleantous, Harold
: APPLICANT: Al-Garawi, Amal
: APPLICANT: Miller, Charles
: APPLICANT: Tomb, Jean-Francois
: APPLICANT: Oomen, Raymond P.
: TITLE OF INVENTION: Identification of Polynucleotides
: TITLE OF INVENTION: Encoding No. US20020115078A1 Helicobacter Polypeptides in the
: FILE REFERENCE: 06133/041002
: CURRENT APPLICATION NUMBER: US/09/881,752A
: CURRENT FILING DATE: 2001-06-15
: PRIOR APPLICATION NUMBER: US 08/833,457
: PRIOR FILING DATE: 1997-04-01
```

```

: NUMBER OF SEQ ID NOS: 370
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 128
: LENGTH: 491
: TYPE: PRF
: ORGANISM: Helicobacter pylori
: FEATURE:
: NAME/KEY: VARIANT
: LOCATION: 84
: OTHER INFORMATION: Xaa = Any Amino Acid
US-09-881-752A-128

Query Match
Best Local Similarity 38.4%; Score 41.5; DB 10; Length 491;
Matches 11; Conservative 1; Mismatches 5; Indels 3; Gaps 1;

QY 2 LNPDMY--LISKDYTGATK 18
DB 173 LNPDMYLGGLTKLDVDRKK 192

RESULT 12
US-09-925-299-1177
: Sequence 1177, Application US/09925299
: Publication No. US20030040617A9
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
: FILE REFERENCE: PA102
: CURRENT APPLICATION NUMBER: US/09/925,299
: CURRENT FILING DATE: 2001-08-10
: PRIOR APPLICATION NUMBER: PCT/US00/05883
: PRIOR FILING DATE: 2000-03-08
: PRIOR APPLICATION NUMBER: 60/124,270
: PRIOR FILING DATE: 1999-03-12
: NUMBER OF SEQ ID NOS: 156
: SOFTWARE: Patent In Ver. 2.0
: SEQ ID NO 1177
: LENGTH: 166
: TYPE: PRF
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: SITE
: LOCATION: (157)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
: NAME/KEY: SITE
: LOCATION: (158)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
: NAME/KEY: SITE
: LOCATION: (163)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-299-1177

Query Match
Best Local Similarity 38.0%; Score 41; DB 9; Length 166;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 ALNPDNYLSKDV 13
DB 141 ALNKNFLISDEL 153

RESULT 13
US-09-925-299-1177
: Sequence 1177, Application US/09925299
: Patent No. US20020055627A1
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
: FILE REFERENCE: PA102
: CURRENT APPLICATION NUMBER: US/09/925,299
: CURRENT FILING DATE: 2001-08-10
: PRIOR APPLICATION NUMBER: PCT/US00/05883
```

;; PRIOR FILING DATE: 2000-03-08
;; PRIOR APPLICATION NUMBER: 60/124,270
;; PRIOR FILING DATE: 1999-03-12
;; NUMBER OF SEQ ID NOS: 1556
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 1177
;; LENGTH: 166
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: SITE
;; LOCATION: (157)
;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;; NAME/KEY: SITE
;; LOCATION: (158)
;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;; NAME/KEY: SITE
;; LOCATION: (163)
;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-299-1177

Query Match 38.0%; Score 41; DB 10; Length 166;
Best Local Similarity 61.5%; Pred. No. 37;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Oy 1 ALNDNYLISKDV 13
Db 141 ALNOKNFIISED 153

RESULT 14
US-09-860-670-117
;; Sequence 117; Application US/09860670
;; Patent No. US20020165137A1
;; GENERAL INFORMATION:
;; APPLICANT: Ruben et al.
;; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;; FILE REFERENCE: PA127P1
;; CURRENT APPLICATION NUMBER: US/09/860,670
;; CURRENT FILING DATE: 2001-05-21
;; Prior application data removed - consult PALM or file wrapper
;; NUMBER OF SEQ ID NOS: 289
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 117
;; LENGTH: 431
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: SITE
;; LOCATION: (205)
;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-860-670-117

Query Match 38.0%; Score 41; DB 9; Length 431;
Best Local Similarity 50.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Oy 6 NYLISKDVTGATKYK 21
Db 388 HTIISASSDGTAKYK 403

RESULT 15
US-09-808-602-60
;; Sequence 60; Application US/09808602
;; Patent No. US20020155115A1
;; GENERAL INFORMATION:
;; APPLICANT: Vernet, Corline A
;; APPLICANT: Fernandes, Elma
;; APPLICANT: Shimkets, Richard A
;; APPLICANT: Heriman, John L
;; APPLICANT: Majumder, Kumud
;; APPLICANT: Mishra, Vishnu

;; APPLICANT: Mezes, Peter S
;; APPLICANT: MacDougall, John
;; TITLE OF INVENTION: No. US20020155115A1el Proteins and Nucleic Acids Encoding Same
;; FILE REFERENCE: 15966-697 CIP
;; CURRENT APPLICATION NUMBER: US/09/808,602
;; PRIOR FILING DATE: 2001-03-14
;; PRIOR APPLICATION NUMBER: 08/800,198
;; PRIOR FILING DATE: 2001-03-05
;; PRIOR APPLICATION NUMBER: 60/186,596
;; PRIOR FILING DATE: 2000-03-03
;; NUMBER OF SEQ ID NOS: 114
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 60
;; LENGTH: 773
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-808-602-60

Query Match 38.0%; Score 41; DB 9; Length 773;
Best Local Similarity 45.5%; Pred. No. 2.3e+02;
Matches 10; Conservative 6; Mismatches 4; Indels 2; Gaps 2;

Oy 2 LMPD-NYLIS-KDVTGATKYK 21
Db 625 VSPDGHVLVSINDYKGLVRYQ 646

Search completed: June 26, 2003, 00:54:45
Job time : 4.44706 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 26, 2003, 00:46:24 ; Search time 3.45882 Seconds
(without alignments)
583.673 Million cell updates/sec

Title: US-09-980-370-8
Perfect score: 108
Sequence: 1 ALNPDNYLISKDVATGATKVKY 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR73:*

1: PIR1:*

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	108	100.0	1526	2	gingipain R (EC 3.4.22.37)
2	108	100.0	1526	2	gingipain R (EC 3.4.22.37)
3	108	100.0	1526	2	gingipain R (EC 3.4.22.37)
4	108	100.0	1526	2	gingipain R (EC 3.4.22.37)
5	108	100.0	1526	2	gingipain R (EC 3.4.22.37)
6	108	100.0	1526	2	gingipain R (EC 3.4.22.37)
7	108	100.0	1526	2	gingipain R (EC 3.4.22.37)
8	108	100.0	1526	2	gingipain R (EC 3.4.22.37)
9	108	100.0	1526	2	gingipain R (EC 3.4.22.37)
10	108	100.0	1526	2	gingipain R (EC 3.4.22.37)
11	108	100.0	1526	2	gingipain R (EC 3.4.22.37)
12	108	100.0	1526	2	gingipain R (EC 3.4.22.37)
13	108	100.0	1526	2	gingipain R (EC 3.4.22.37)
14	108	100.0	1526	2	gingipain R (EC 3.4.22.37)
15	108	100.0	1526	2	gingipain R (EC 3.4.22.37)
16	108	100.0	1526	2	gingipain R (EC 3.4.22.37)
17	108	100.0	1526	2	gingipain R (EC 3.4.22.37)
18	108	100.0	1526	2	gingipain R (EC 3.4.22.37)
19	108	100.0	1526	2	gingipain R (EC 3.4.22.37)
20	108	100.0	1526	2	gingipain R (EC 3.4.22.37)
21	108	100.0	1526	2	gingipain R (EC 3.4.22.37)
22	108	100.0	1526	2	gingipain R (EC 3.4.22.37)
23	108	100.0	1526	2	gingipain R (EC 3.4.22.37)
24	108	100.0	1526	2	gingipain R (EC 3.4.22.37)
25	108	100.0	1526	2	gingipain R (EC 3.4.22.37)
26	108	100.0	1526	2	gingipain R (EC 3.4.22.37)
27	108	100.0	1526	2	gingipain R (EC 3.4.22.37)
28	108	100.0	1526	2	gingipain R (EC 3.4.22.37)
29	108	100.0	1526	2	gingipain R (EC 3.4.22.37)

30	42.5	39.4	1965	2	T33216	hypothetical prote
31	42	38.9	169	2	C81192	hypothetical prote
32	42	38.9	232	2	P97914	conserved hypotet
33	42	38.9	310	2	B97777	chlorodioxin-disulf
34	42	38.9	318	2	D86325	hypothetical prote
35	42	38.9	383	2	S38904	hypothetical prote
36	42	38.9	528	2	H70163	phenylalanine-LNA
37	42	38.9	561	2	T34368	hypothetical prote
38	42	38.9	653	2	T01274	hypothetical prote
39	42	38.9	983	1	VCLJVS	env polyprotein pr
40	42	38.9	983	1	E45390	hemolysin A toxin
41	42	38.9	998	2	T00227	myosin IXA [import
42	42	38.9	1777	2	T34369	cytochrome P450-11
43	42	38.9	2548	2	E59435	hypothetical prote
44	41.5	38.4	512	2	T48462	hypothetical prote
45	41.5	38.4	704	2	D97723	hypothetical prote

ALIGNMENTS

RESULT 1
S49763
gingipain R (EC 3.4.22.37) precursor - Porphyromonas gingivalis (fragment)
C:Species: Porphyromonas gingivalis
C:Date: 05-Mar-1995 #sequence_revision 12-May-1995 #text_change 31-Mar-1997
C:Accession: S49763
R:Aduse-Opoku, J.; Mull, J.; Slaney, J.M.; Rangarajan, M.; Curtis, M.A.
submitted to the EMBL Data Library, November 1994
A:Description: Cloning, sequence analysis and expression in Escherichia coli of prpRI
A:Reference number: S49763
A:Accession: S49763
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1526 <ADU>
A:Cross-references: EMBL:X82680
A:Genetics:
A:Gene: prpRI
C:Keywords: cysteine proteinase; hydrolase

Query Match 100.0% Score 108; DB 2; Length 1526;
Best Local Similarity 100.0%; Pred. No. 8 Re-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNPDNYLISKDVATGATKVKY 21
DB 1198 ALNPDNYLISKDVATGATKVKY 1218

RESULT 2
A53426
gingipain R (EC 3.4.22.37) precursor - Porphyromonas gingivalis
N:Alternate names: 50K high molecular mass arginine-specific cysteine proteinase; HGP
C:Species: Porphyromonas gingivalis
C:Date: 10-Feb-1995 #sequence_revision 10-Feb-1995 #text_change 08-Oct-1999
C:Accession: A53426; D53113
R:Payloft, N.; Potempa, J.; Pike, R.N.; Prochazka, V.; Kiefer, M.C.; Travis, J.; Barr
J.; Biol. Chem. 270, 1007-1010, 1995
A:Title: Molecular cloning and structural characterization of the Arg-gingipain prote
A:Reference number: A53426; MUID:95138080; PMID:7636351
A:Accession: A53426
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1704 <PAV>
A:Cross-references: GB:U15282; NID:9557067; PIDD:AAA69539.1; PID:9557068
R:Pike, R.; McGraw, W.; Potempa, J.; Travis, J.
J. Biol. Chem. 269, 406-411, 1994
A:Title: Lysine- and arginine-specific proteinases from Porphyromonas gingivalis. Iso
A:Accession: D53113
A:Reference number: A53113; MUID:94103245; PMID:8276827
A:Status: preliminary
A:Molecule type: protein
A:Residues: 228-249 <PIK>

A:Experimental source: H66
 A:Note: sequence extracted from NCBI backbone (NCBIP:141694)
 C:Keywords: cysteine proteinase; hydrolase

Query Match 100.0%; Score 108; DB 2; Length 1704;
 Best Local Similarity 100.0%; Pred. No. 9,9e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ALNPDNYLSKDVGTGATKVKY 21
 |||
 Db 1196 ALNPDNYLSKDVGTGATKVKY 1216

RESULT 3
 T30836
 Lysine-specific cysteine proteinase porphyrain (EC 3.4.22.-) - Porphyromonas gingivalis
 N:Alternate names: lysine-specific cysteine proteinase I, 60k
 C:Species: Porphyromonas gingivalis
 C>Date: 22-Oct-1999 #sequence-revision 22-Oct-1999 #text-change 17-Nov-2000
 C:Accession: T30836; T30837; T30526; A53113
 R:Barkocy-Gallagher, G.A.; Han, N.; Patel, J.M.; Whitlock, J.; Prognulske-Fox, A.; Lantz, J.; Bacteriol. 178:2734-2741, 1996
 A:Title: Analysis of the prt gene encoding porphyrain, a cysteine proteinase of Porphyromonas gingivalis
 A:Reference number: Z20895; MUID:96213011; PMID:8531659
 A:Accession: T30836
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1732 <HAB>
 R:Stakesli, N.; Chel, S.M.; Reynolds, E.C.
 A:Cross-references: EMBL:U42210; NID:91314325; PID:91314326; PIDN:AA06565.1
 submitted to the EMBL Data Library, October 1996
 A:Reference number: Z20896
 A:Accession: T30837
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-795; T: 797-1389; 'N', 1391-1478; 'Y', 1480-1732 <LEH>
 R:Lewis, J.P.; Macrina, F.L.
 A:Cross-references: EMBL:U73566; NID:92182811; FID:92182812; FIDN:AA06809.1
 Infect. Immun. 65: 3035-3042, 1998
 A:Title: IS195, an insertion sequence-like element associated with protease genes in Porphyromonas gingivalis
 A:Reference number: Z20844; MUID:9828016; PMID:9532553
 A:Accession: T30526
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1350; 'N', 1352-1363; 'Y', 1365-1447; 'H', 1449-1732 <LEH>
 R:Plake, R.; McGraw, W.; Potempa, J.; Travis, J.
 J. Biol. Chem. 269: 406-411, 1994
 A:Title: Lysine- and arginine-specific proteinases from Porphyromonas gingivalis. Isolation and characterization of the genes
 A:Reference number: A53113; MUID:94103245; PMID:8276827
 A:Accession: A53113
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 228-249 <PIK>
 A:Experimental source: H66
 A:Note: sequence extracted from NCBI backbone (NCBIP:141690)
 C:Genetics:
 A:Gene: prt; prt
 C:Keywords: cysteine proteinase; hydrolase

Query Match 100.0%; Score 108; DB 2; Length 1732;
 Best Local Similarity 100.0%; Pred. No. 1e-08;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ALNPDNYLSKDVGTGATKVKY 21
 |||
 Db 1216 ALNPDNYLSKDVGTGATKVKY 1236

RESULT 4
 T28651
 hemagglutinin A - Porphyromonas gingivalis
 C:Species: Porphyromonas gingivalis

C>Date: 15-Oct-1999 #sequence-revision 15-Oct-1999 #text-change 03-Aug-2001
 C:Accession: T28651
 R:Han, N.; Whitlock, J.; Prognulske-Fox, A.
 Infect. Immun. 64: 4000-4007, 1996
 A:Title: The hemagglutinin gene A (haga) of Porphyromonas gingivalis 381 contains four copies of a 120-bp repeat
 A:Reference number: Z20494; MUID:97047672; PMID:8926051
 A:Accession: T28651
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-2628 <HAN>
 A:Cross-references: EMBL:U41807; NID:91552410; PID:91469916; PIDN:AA017128.1
 C:Genetics:
 A:Gene: haga

Query Match 100.0%; Score 108; DB 2; Length 2628;
 Best Local Similarity 100.0%; Pred. No. 1.6e-08;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ALNPDNYLSKDVGTGATKVKY 21
 |||
 Db 744 ALNPDNYLSKDVGTGATKVKY 764

RESULT 5
 T27658
 hypothetical protein ZK1037.3 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 15-Oct-1999 #sequence-revision 15-Oct-1999 #text-change 15-Oct-1999
 C:Accession: T27658
 R:Basham, V.
 submitted to the EMBL Data Library, October 1996
 A:Reference number: Z20401
 A:Accession: T27658
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-334 <MWL>
 A:Cross-references: EMBL:Z81142; PIDN:CA03502.1; GSPDB:G000023; CESP:ZK1037.3
 A:Experimental source: clone ZK1037
 C:Genetics:
 A:Gene: CESP:ZK1037.3
 A:Map position: 5
 A:Introns: 57/3; 199/3; 248/3; 292/1

Query Match 50.9%; Score 55; DB 2; Length 334;
 Best Local Similarity 64.7%; Pred. No. 0.56;
 Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Oy 2 LNPDPYLSKDVGTGATK 18
 |||
 Db 22 LNPDPYLSKDVGTGATK 38

RESULT 6
 F81042
 hemolysin activation protein HecB, probable NMB1780 [imported] - Neisseria meningitidis
 C:Species: Neisseria meningitidis
 C>Date: 31-Mar-2000 #sequence-revision 31-Mar-2000 #text-change 19-Jan-2001
 C:Accession: F81042
 R:Reichlin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, R.; Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.; Li, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignan, V.; Piazza, M.
 Science 287: 1809-1815, 2000
 A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappelli, R.; Venter, A.; et al.
 A:Reference number: AB1000; MUID:20157555; PMID:10710307
 A:Accession: F81042
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-580 <PEY>
 A:Cross-references: GB:AE002528; GB:AE002098; NID:97227034; PIDN:AAFA2120.1; PID:9722
 A:Experimental source: serogroup B, strain MC58
 C:Genetics:
 A:Gene: NMB1780

Query Match 46.3%; Score 50; DB 2; Length 580;
 Best Local Similarity 45.0%; Pred. No. 6.5;
 Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

OY 2 LMPDNYLSKDYTGATKVKY 21
 DB 56 LKNNYVLSDEPTCTRVNY 75

RESULT 7

AB1989 Probable periplasmic protein NMA0687 [Imported] - Neisseria meningitidis (strain 22491)

C:Species: Neisseria meningitidis
 C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
 C:Accession: AB1989
 R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel, J.; Holtroyd, S.; Jagsels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream, Nature 404, 502-506, 2000
 A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis 22491.
 A:Reference number: AB1773; MUID:2022356; PMID:10761919
 A:Accession: AB1989
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-580 <PAR>
 A:Cross-references: GB:AL162753; GB:AL157959; NID:g7379120; PIDN:CAB83973.1; PID:g737941
 A:Experimental source: serogroup A, strain 22491
 C:Genetics: NMA0686; NMA0687

Query Match 46.3%; Score 50; DB 2; Length 580;
 Best Local Similarity 45.0%; Pred. No. 6.5;
 Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

OY 2 LMPDNYLSKDYTGATKVKY 21
 DB 56 LKNNYVLSDEPTCTRVNY 75

RESULT 8

S40209 tubulin gamma chain - fungus (Cochliobolus heterostrophus)

C:Species: Cochliobolus heterostrophus, Bipolaris maydis
 C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 13-Aug-1999
 C:Accession: S40209
 R:Parkinson, C.; Luo, H.; Knight, A.; Ahlquist, J.; Perlman, M.H.
 A:Submitted to the EMBL Data Library, August 1993
 A:Description: Phylogenetic analyses using the gamma tubulin gene.
 A:Reference number: S40209
 A:Accession: S40209
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-265 <PAR>
 A:Cross-references: EMBL:X74455; NID:g437988; PIDN:CAA52464.1; PID:g437989
 C:Genetics:
 A:Introns: 136/3
 C:Superfamily: tubulin

Query Match 45.4%; Score 49; DB 2; Length 265;
 Best Local Similarity 64.3%; Pred. No. 4;
 Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 3 LMPDNYLSKDYTGATG 16
 DB 22 NPENFYIHHDTGTA 35

RESULT 9

B46108 Outer capsid protein VP7 - human rotavirus (strain 1321)

C:Species: human rotavirus
 C:Date: 07-Apr-1994 #sequence_revision 25-Apr-1997 #text_change 20-Sep-1999
 C:Accession: B46108

R:Das, M.; Dunn, S.J.; Woode, G.N.; Greenberg, H.B.; Rao, C.D.

Virology 194, 374-379, 1993

A:Title: Both surface proteins (VP4 and VP7) of an asymptomatic neonatal rotavirus st

A:Reference number: A46108; MUID:93242771; PMID:8386881

A:Contents: 1321, serotype 10

A:Accession: B46108

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-326 <DAS>

A:Cross-references: GB:L07658; NID:g310326; PIDN:AAA47308.1; PID:g310327

A:Note: sequence extracted from NCBI Backbone (NCBI:130701, NCBI:130693)

A:Note: the sequence in Genbank entry R01VP7B, release 109, (PID:g310325) has the cod

C:Superfamily: rotavirus glycoprotein VP7

Query Match 42.6%; Score 46; DB 2; Length 326;
 Best Local Similarity 40.0%; Pred. No. 15;
 Matches 8; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

OY 2 LMPDNYLSKDYTGATKVKY 21
 DB 48 VNPQVGINLPITGSDMSY 67

RESULT 10

S31572

cinamyl-alcohol dehydrogenase (EC 1.1.1.15) - alfalfa

C:Species: Medicago sativa (alfalfa)

C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 15-Oct-1999

C:Accession: S31572

R:van Doorselaere, J.; van Montagu, M.; Inze, D.

A:Submitted to the EMBL Data Library, January 1993

A:Description: Isolation and characterization of poplar and alfalfa cinamyl alcohol

A:Reference number: S31571

A:Accession: S31572

A:Molecule type: mRNA

A:Residues: 1-358 <VAN>

A:Cross-references: EMBL:219573; NID:g19594; PIDN:CAA79625.1; PID:g19595

C:Superfamily: alcohol dehydrogenase; long-chain alcohol dehydrogenase homology

C:Keywords: alcohol metabolism; metalloprotein; oxidoreductase; zinc

F:33-341/Domain: long-chain alcohol dehydrogenase homology <LADH>

F:48,70,164/Binding site: zinc, catalytic (Cys, His, Cys) #status predicted

Query Match 41.7%; Score 45; DB 2; Length 358;
 Best Local Similarity 57.1%; Pred. No. 25;
 Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 2 LMPDNYLSKDYTG 15
 DB 225 LGADNYLVSDIVG 238

RESULT 11

H70209

decorin binding protein A (dbpa) homolog - Lyme disease spirochete plasmid A/1p54

C:Species: Borrelia burgdorferi (Lyme disease spirochete)

C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 08-Oct-1999

C:Accession: H70209

R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Tadhigra, R.; Wh

son, D.; Peterson, J.; Kierlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vu

lman, C.; Garland, S.; Fujil, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.

Nature 390, 580-586, 1997

A:Authors: Smith, H.O.; Venier, J.C.

A:Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.

A:Reference number: A70100; MUID:98065943; PMID:9403685

A:Accession: H70209

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-191 <KLE>

A:Cross-references: GB:AE000790; NID:g2690224; PIDN:AC66250.1; PID:g2690249; TIGR:BB

A:Experimental source: strain B31

C:Genetics:

A:Genome: plasmid

C:Superfamily: Lyme disease spirochete plasmid probable decorin binding protein A

Query Match 40.7%: Score 44: DB 2: Length 191:
 Best Local Similarity 60.0%: Pred. No. 18:
 Matches 9: Conservative 3: Mismatches 3: Indels 0: Gaps 0:

OY 6 NYLISKDVTGATKVK 20
 DB 20 NLISCGTATKIR 34

RESULT 12

B96533
 Hypothetical protein F14J22.12 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C:Accession: B96533

R:Metadlogis: A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;

ansen, N.F.; Hughes, B.; Hultzer, L.

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: B96533

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-315 <STO>

A:Cross-references: GB:AE005173; NID:g10120425; PIDN:AA613050.1; GSPDB:GN00141

C:Genetics:

A:Gene: F14J22.12

A:Map position: 1

Query Match

Best Local Similarity 61.5%: Score 44: DB 2: Length 315:
 Matches 8: Conservative 3: Mismatches 2: Indels 0: Gaps 0:

OY 1 ALPNPNYLSKDV 13
 DB 34 SLNPONDVYSKDV 46

RESULT 13

AB1656
 cytochrome beta-lyase homolog lln1787 [imported] - Listeria innocua (strain Clp11262

C:Species: Listeria innocua

C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001

C:Accession: AB1656

R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bioecker,

.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entlian, K.D.; Fsihi, H.

A:Authors: Krefl, J.; Kuhn, M.; Kunst, F.; Kurapkak, G.; Madueno, E.; Maltournam, A.; Ma

Ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,

A:Title: Comparative genomics of Listeria species

A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: AB1656

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-390 <GLA>

A:Cross-references: GB:AL592022; PIDN:CAC97018.1; PID:g16414274; GSPDB:GN00178

A:Experimental source: strain Clp11262

C:Genetics:

A:Gene: lln1787

C:Superfamily: O-succinylhomoserine (thiol)-lyase

Query Match 40.7%: Score 44: DB 2: Length 390:
 Best Local Similarity 61.5%: Pred. No. 39:
 Matches 8: Conservative 3: Mismatches 2: Indels 0: Gaps 0:

OY 5 DNTLISKDVTGAT 17
 DB 93 DHTLISKDVTGAT 105

RESULT 14

B84553
 Probable protein kinase/endonuclease [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001

C:Accession: B84553

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanden, S.E.; Umayam, L.; Tallon,

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: B84553

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-393 <STO>

A:Cross-references: GB:AE002093; NID:g4914373; PIDN:AA032909.1; GSPDB:GN00139

C:Genetics:

A:Gene: AT2G17520

A:Map position: 2

Query Match

Best Local Similarity 56.2%: Score 44: DB 2: Length 393:
 Matches 9: Conservative 3: Mismatches 4: Indels 0: Gaps 0:

OY 4 PDNYLSKDVYATKV 19
 DB 125 PDNYLSKDVYATKV 140

RESULT 15

A86845
 Hypothetical protein yscA [imported] - Lactococcus lactis subsp. lactis (strain IL140

C:Species: Lactococcus lactis subsp. lactis

C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001

C:Accession: A86845

R:Boletín, A.; Wincker, P.; Manger, S.; Tallon, O.; Malarme, K.; Weissenbach, J.; Eh

Genome Res. 11, 731-753, 2001

A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis

A:Reference number: A86625; MUID:21235186; PMID:1137471

A:Accession: A86845

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-659 <STO>

A:Cross-references: GB:AE005176; PID:g12724783; PIDN:AAK05859.1; GSPDB:GN00146

A:Experimental source: strain IL1403

C:Genetics:

A:Gene: yscA

Query Match 40.7%: Score 44: DB 2: Length 639:
 Best Local Similarity 81.8%: Pred. No. 67:
 Matches 9: Conservative 0: Mismatches 2: Indels 0: Gaps 0:

OY 4 PDNYLSKDVYATKV 14
 DB 452 PDNYLSKDVYATKV 462

Search completed: June 26, 2003, 00:52:22

Job time : 3.45882 secs

Thu Jun 26 11:58:21 2003

us-09-980-370-8.rsp

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 26, 2003, 00:44:19 ; Search time 2.22353 Seconds

(without alignments)
391.721 Million cell updates/sec

Title: US-09-980-370-8

Perfect score: 108

Sequence: 1 ALNDNDYLSKDVGTATKVKY 21

Scoring table: GAPOP 10.0 , GAPEXT 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	108	100.0	989	1	PRTH_PORGI
2	108	100.0	2628	1	HAGA_PORGI
3	49	45.4	265	1	TBG_COCHI
4	45	41.7	358	1	CADH_MEDSA
5	45	41.7	418	1	LEU2_THETN
6	45	41.7	462	1	TBG2_EUPCR
7	44	40.7	191	1	DBPA_BORBU
8	44	40.7	789	1	CAD2_HUMAN
9	43.5	40.3	313	1	GAS2_HUMAN
10	43	39.8	209	1	KGUA_CLOAB
11	43	39.8	364	1	CARA_LACPL
12	43	39.8	369	1	METB_HAEIN
13	43	39.8	1588	1	AKO1_YEAST
14	43	39.8	2512	1	FOLN_STNDV
15	43	39.8	2514	1	FOLN_STNDV
16	42.5	39.4	311	1	O2T1_HUMAN
17	42.5	39.4	314	1	GAS2_MOUSE
18	42.5	39.4	851	1	MCEL_FOWPV
19	42	38.9	310	1	TRXB_RICCN
20	42	38.9	528	1	SYFA_BORBU
21	42	38.9	588	1	SYD_PASMU
22	42	38.9	982	1	ENV_VILV
23	42	38.9	983	1	ENV_VILV
24	41	38.0	302	1	LGUL_SCHPO
25	41	38.0	417	1	LE22_THEMA
26	41	38.0	618	1	CHAA_CLOBI
27	41	38.0	710	1	PKNG_MYXPA
28	41	38.0	808	1	SECA_MYXPA
29	41	38.0	1607	1	LMG1_MOUSE
30	41	38.0	1609	1	LMG1_MOUSE
31	40.5	37.5	524	1	LCYE_HUMAN
32	40.5	37.5	817	1	PSPA_PYRAB
33	40.5	37.5	819	1	PSPA_PYRAB

34	40.5	37.5	821	1	PSPA_PYRAB
35	40	37.0	278	1	PR28_MYCLE
36	40	37.0	297	1	EX53_MYCPU
37	40	37.0	354	1	MTD2_HERAU
38	40	37.0	446	1	TBG_SCHRO
39	40	37.0	452	1	TBG_SCHRO
40	40	37.0	468	1	TBG_PLAHO
41	40	37.0	532	1	RRPO_BPCA
42	40	37.0	586	1	VLGN_BEPV
43	40	37.0	942	1	AMPN_MANSE
44	40	37.0	1786	1	LMBI_HUMAN
45	40	37.0	1786	1	LMBI_MOUSE

ALIGNMENTS

RESULT 1
ID PRTH_PORGI STANDARD: PRT: 989 AA.
AC P46071.1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Protease prth (EC 3.4.22.-).
GN PRTH.
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; Bacteroidetes; Bacteroides; Bacteroidales;
OC Porphyromonadaceae; Porphyromonas.
OX NCBI_Taxid=837;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=95012612; PubMed=7927685;
RX STRAIN=WB3;
RA Fletcher H.M., Schenkein H.A., Macrina F.L.;
RT "Cloning and characterization of a new protease gene (prth) from
KT Porphyromonas gingivalis."
RL Infect. Immun. 62:4279-4286(1994).
RN [2]
RP ERRATUM.
RA Fletcher H.M., Schenkein H.A., Macrina F.L.;
RL Infect. Immun. 62:5707-5707(1994).
CC -!- FUNCTION: CLEAVES HUMAN COMPLEMENT COMPONENT C3. MAY ENABLE
CC P. GINGIVALIS TO EVADE COMPLEMENT-MEDIATED KILLING DURING THE
CC IMMUNE RESPONSE. PLAYS AN IMPORTANT ROLE IN SOFT TISSUE INFECTIONS
CC AND IS A VIRULENCE FACTOR.
CC -!- SUBCELLULAR LOCATION: IN MEMBRANE VESICLES.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C25.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb.slb.ch/announce/>
CC or send an email to license@isb.slb.ch).
CC EMBL: L27483; AA51298.1;
CC HSSP: P23882; 1PMT.
CC DR INTERPRO: IPR001769; Peptidase_C25.
CC DR InterPro: IPR002376; formyl_transf.
CC DR Pfam: PF01364; Peptidase_C25; 1.
CC FT Hydrolyase; Thiol protease; Repeat; Virulence.
CC REPEAT 270 323
CC FT REPEAT 528 581
CC SEQUENCE 989 AA: 110238 MW: FA85FEBA3AC8944C CRC64:
Query Match 100.0%; Score 108; DB 1; Length 989;
Best Local Similarity 100.0%; Pred. No. 2.3e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```

CC -1- CATALYTIC ACTIVITY: Cinnamyl alcohol + NADP(+) = cinnamaldehyde +
CC NADPH.
CC -1- COFACTOR: ZINC (BY SIMILARITY).
CC -1- PATHWAY: Lignin synthesis
CC -1- TISSUE SPECIFICITY: MOST ACTIVELY EXPRESSED IN STEM, HYPOCOTYL AND
CC ROOT TISSUE.
CC -1- SIMILARITY: BELONGS TO THE ZINC-CONTAINING ALCOHOL DEHYDROGENASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: 219573; CAI79625.1; -
CC DR EMBL: AF083332; AAC35845.1; -
CC DR PIR: S31572; S31572.
CC DR InterPro: IPR002328; ADH_Zinc.
CC DR InterPro: IPR002085; Adh_zn_family.
CC DR Pfam: PF00107; adh_zinc; 1.
CC DR PROSITE: PS00059; ADH_ZINC; 1.
CC KW Oxidoreductase; NADP; zinc; Lignin biosynthesis.
CC FT METAL 48 48 ZINC (CATALYTIC) (BY SIMILARITY).
CC FT METAL 70 70 ZINC (CATALYTIC) (BY SIMILARITY).
CC FT METAL 101 101 ZINC (SECOND ATOM) (BY SIMILARITY).
CC FT METAL 104 104 ZINC (SECOND ATOM) (BY SIMILARITY).
CC FT METAL 107 107 ZINC (SECOND ATOM) (BY SIMILARITY).
CC FT METAL 115 115 ZINC (SECOND ATOM) (BY SIMILARITY).
CC FT METAL 164 164 ZINC (CATALYTIC) (BY SIMILARITY).
CC SQ SEQUENCE 358 AA; 38948 MW; FBA609408D01BF56 CRC64;

Query Match 41.7%; Score 45; DB 1; Length 358;
Best Local Similarity 57.1%; Pred. No. 9.9;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 LAMPDNYLSKDVGTG 15
DB 225 LGADNYLVSDTVG 238

RESULT 5
LEU2_THETN
ID LEU2_THETN STANDARD: PRT: 418 AA.
AC QBRDK2:
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 3-Isopropylmalate dehydratase large subunit (EC 4.2.1.33)
DE (Isopropylmalate isomerase) (Alpha-IPM isomerase) (IPMI).
GN LEUC OR TYP0017
OS Thermoanaerobacter tengcongensis.
OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;
OC Thermoanaerobacteriaceae; Thermoanaerobacter.
OX NCBI_TaxID=119072;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MB4 / JCM 11007;
RX MEDLINE=21992816; PubMed=11997336;
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
RA Tan H., Chen R., Wang J., Yu J., Yang H.;
RA "A complete sequence of T. tengcongensis genome.";
RT Genome Res. 12:689-700(2002).
DE -1- FUNCTION: Catalyzes the isomerization between 2-isopropylmalate
DE and 3-isopropylmalate, via the formation of 2-isopropylmalate
CC -1- CATALYTIC ACTIVITY: 3-Isopropylmalate = 2-Isopropylmalate +
CC H(2)O.
CC -1- CATALYTIC ACTIVITY: 2-Isopropylmalate + H(2)O = 2-
CC Isopropylmalate.
CC -1- PATHWAY: Leucine biosynthesis; second step.

```

```

CC -1- SUBUNIT: Heterodimer of leuc and leud (by similarity).
CC -1- SIMILARITY: BELONGS TO THE ACONITASE/IPM ISOMERASE FAMILY. LEUC 2
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE012976; AAM23334.1; -
CC DR EMBL: AF001030; Aconitase_N.
CC DR Pfam: PF00330; aconitase; 1.
CC DR ProDom: PD000511; Aconitase_N; 1.
CC DR PROSITE: PS00450; ACONITASE_1; 1.
CC DR PROSITE: PS01244; ACONITASE_2; FALSE_NEG.
CC KW Leucine biosynthesis; lyase; iron-sulfur; 4Fe-4S; Complete proteome.
CC FT METAL 298 298 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
CC FT METAL 358 358 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
CC FT METAL 361 361 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
CC SQ SEQUENCE 418 AA; 45510 MW; AC3DE0D6E7A97AE CRC64;

Query Match 41.7%; Score 45; DB 1; Length 418;
Best Local Similarity 47.4%; Pred. No. 12;
Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 LAMPDNYLSKDVGTGATV 19
DB 60 ALVDPHFVPSKDIKSAEQV 78

RESULT 6
TBG2_EUPCR
ID TBG2_EUPCR STANDARD: PRT: 462 AA.
AC PS4404:
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tubulin gamma-2 chain (Gamma-2 tubulin).
DE Eukaryotes crassus.
OC Eukaryota; Alveolata; Ciliophora; Spirotrichea; Hypotrichia;
OC Euplotida; Euplotidae; Moneuplotes.
OX NCBI_TaxID=5936;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98192540; PubMed=9524221;
RA Tan M., Heckmann K.;
RA "The two gamma-tubulin-encoding genes of the ciliate Euplotes crassus
RA differ in their sequences, codon usage, transcription initiation
RA sites and poly(A) addition sites.";
RL Gene 210:53-60(1998).
CC -1- FUNCTION: TUBULIN IS THE MAJOR CONSTITUENT OF MICROTUBULES.
CC SUCH AS THE SPINDLE POLES OR THE CENTROSOME, SUGGESTING THAT IT
CC IS INVOLVED IN THE MINUS-END NUCLEATION OF MICROTUBULE ASSEMBLY.
CC -1- SIMILARITY: BELONGS TO THE TUBULIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X85235; CAI58490.1; -
CC DR EMBL: Y09551; CAI70742.1; -
CC DR InterPro: IPR000217; Tubulin.
CC DR InterPro: IPR003008; Tubulin_ftsz.
CC DR Pfam: PF00091; tubulin; 1.
CC DR PRINTS: PR01161; TUBULIN.

```

DR PROSITE: P500227: TUNRLIN; 1.
 KM Microtubules: GTP-binding; Multigene family.
 FT NP_BIND 148 GTP (POTENTIAL).
 SO SEQUENCE 462 AA: 52117 MW: FF65B755E9AC846 CRC64.
 Query Match 41.7%; Score 45; DB 1; Length 462;
 Best Local Similarity 50.0%; Pred. No. 13;
 Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
 OY 3 NP0NLYSKDVGCA 16
 DB 87 NPEFVSMTMDGA 100
 RESULT 7
 DBPA_BORBU STANDARD; PRT: 191 AA.
 AC 050917: Q9R805: 006876;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Decorin binding protein A precursor.
 GN DBPA OR BBA24.
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OC Plasmid p954.
 OX Bacteria: Spirochaetales; Spirochaetaceae; Borrelia.
 RN NCBI_TaxID=139;
 RP SEQUENCE FROM N.A.
 RC STRAIN=297;
 RA MEDLINE=95369901: PubMed=7642279;
 RT Guo B.P., Norris S.J., Rosenberg L.C., Hook M.;
 RT "Adherence of Borrelia burgdorferi to the proteoglycan decorin.";
 RT Infect. Immun. 63:3467-3472(1995).
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 35210 / B31, 3028, HBNC, LP7, and SH-2-82;
 RA MEDLINE=98065943: PubMed=9784533;
 RA Roberts W.C., Mullikin B.A., Lathigra R., Hanson M.S.;
 RT "Molecular analysis of sequence heterogeneity among genes encoding
 RT decorin binding proteins A and B of Borrelia burgdorferi sensu lato.";
 RT Infect. Immun. 66:5275-5285(1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 35210 / B31;
 RA MEDLINE=98065943: PubMed=9403685;
 RA Fraser C.M., Castjens S., Huang W.M., Clayton R.A.,
 RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.R., Gwinn M.,
 RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D., Hanson M.,
 RA Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S.,
 van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Meldrum J.,
 Utechtack T., Mather L., McDonald L., Atliach P., Bowman C.,
 Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
 Smith H.O., Venter J.C.;
 RT "Genomic sequence of a Lyme disease spirochete, Borrelia
 RT burgdorferi.";
 RT Nature 390:580-586(1997).
 RL Nature 390:580-586(1997).
 CC -1- FUNCTION: Binds to decorin which may mediate the adherence of
 CC B. burgdorferi to collagen fibers in skin and other tissues.
 CC -1- SIMILARITY: BELONGS TO THE DECORIN-BINDING PROTEIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U75866; AAD05353.1;
 DR EMBL: AF069253; AAC70026.1;
 DR EMBL: AF069255; AAC70030.1;
 DR EMBL: AF069257; AAC70034.1;

DR EMBL: AF069286; AAC70064.1;
 DR EMBL: AF069269; AAC70047.1;
 DR EMBL: AF069275; AAC70053.1;
 DR EMBL: AF069290; AAC66250.1;
 DR TIGR: BBA24;
 DR InterPro: IPR003332: Decorin_bind.
 DR Pfam: PF02352: Decorin_bind.
 KW Signal; Plasmid; Complete proteome.
 FT SIGNAL 1 29
 FT CHAIN 30 191
 FT VARIANT 34 34
 FT VARIANT 37 37
 FT VARIANT 41 43
 FT VARIANT 52 52
 FT VARIANT 55 56
 FT VARIANT 128 128
 FT VARIANT 140 140
 FT VARIANT 169 169
 FT VARIANT 172 173
 FT VARIANT 179 191
 SO SEQUENCE 191 AA: 21213 MW: ADIFL9IC774AFEA1 CRC64;
 Query Match 40.7%; Score 44; DB 1; Length 191;
 Best Local Similarity 60.0%; Pred. No. 7.3;
 Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 OY 6 NYLISKDVGATKVK 20
 DB 20 NLLISGLTGATKIR 34
 RESULT 8
 CAD9_HUMAN STANDARD; PRT: 789 AA.
 ID CAD9_HUMAN
 AC G9ULB4;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Cadherin-9 precursor.
 GN CDH9.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC PubMed=10861224;
 RA Shimoyama Y., Tsujimoto G., Kitajima M., Natori M.;
 RT "Identification of three human type-II classic cadherins and frequent
 RT heterophilic interactions between different subclasses of type-II
 RT classic cadherins.";
 RT Biochem. J. 349:159-167(2000).
 RL Biochem. J. 349:159-167(2000).
 CC -1- FUNCTION: CADHERINS ARE CALCIUM DEPENDENT CELL ADHESION PROTEINS.
 CC THEY PREFERENTIALLY INTERACT WITH THEMSELVES IN A HOMOPHILIC
 CC MANNER IN CONNECTING CELLS. CADHERINS MAY THUS CONTRIBUTE TO THE
 CC SORTING OF HETEROGENEOUS CELL TYPES.
 CC -1- SIMILARITY: CONTAINS 5 CADHERIN DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AB035302; BAA87416.1;
 DR HSSP: P15116; INCT.

DEVELOPMENTAL STAGE: SPECIFICALLY EXPRESSED AT GROWTH ARREST

use by non-profit institutions as long as its content is in no way


```

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-84148439; PubMed-6322438;
RA Strauss E.G., Rice C.M., Strauss J.H.;
RT "Complete nucleotide sequence of the genomic RNA of Sindbis virus.";
RL Virology 133:92-110(1984).
RN [2]
RP SEQUENCE OF 1-54 FROM N.A.
RX MEDLINE-83268700; PubMed-6308269;
RA Ou J.H., Strauss E.G., Strauss J.H.;
RT "The 5'-terminal sequences of the genomic RNAs of several
  alphaviruses.";
RL J. Mol. Biol. 168:1-15(1983).
RN [3]
RP SEQUENCE OF 1429-2512 FROM N.A.
RX MEDLINE-83299955; PubMed-6557423;
RA Strauss E.G., Rice C.M., Strauss J.H.;
RT "Sequence coding for the alphavirus nonstructural proteins is
  interrupted by an opal termination codon.";
RL Proc. Natl. Acad. Sci. U.S.A. 80:5271-5275(1983).
RN [4]
RP SEQUENCE OF 2431-2512 FROM N.A.
RX MEDLINE-83039346; PubMed-6291034;
RA Ou J.H., Rice C.M., Dalgarno L., Strauss E.G., Strauss J.H.;
RT "Sequence studies of several alphavirus genomic RNAs in the region
  containing the start of the subgenomic RNA.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:5235-5239(1982).
CC -1- FUNCTION: NSP2 MAY BE INVOLVED IN RNA BINDING DURING REPLICATION.
CC -1- PFM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- MISCELLANEOUS: READTHROUGH OF THE TERMINATOR CODON UGA OCCURS
  BETWEEN THE CODONS FOR 1896-TYR AND 1897-LEU.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
CC EMBL: J02363; AAA6975.1; ALT_SEQ.
CC PIR: A03917; MNWVS.
DR DR
DR MEROPS: C09.001;
DR InterPro: IPR002589; Alpp.
DR InterPro: IPR002620; Peptidase_C9.
DR InterPro: IPR001788; RNA_dep_RNAPol2.
DR InterPro: IPR000606; Viral_helicase1.
DR pfam: PF00978; RNA_dep_RNAPol2; 1.
DR pfam: PF01443; Viral_helicase1; 1.
DR pfam: PF01661; Alpp; 1.
DR pfam: PF01707; Peptidase_C9; 1.
DR SMART: SM00506; Alpp; 1.
KW Polyprotein; Nonstructural protein; RNA-binding; Helicase.
FT CHAIN 1 540 NONSTRUCTURAL PROTEIN NSP2.
FT CHAIN 541 1347 NONSTRUCTURAL PROTEIN NSP2.
FT CHAIN 1348 1896 NONSTRUCTURAL PROTEIN NSP3.
FT CHAIN 1897 2512 NONSTRUCTURAL PROTEIN NSP4.
FT CHAIN 2513 2512 NONSTRUCTURAL PROTEIN NSP4.
SO SEQUENCE 2512 AA; 279546 MW; F3656FCBB8495726 CRC64;

Query Match 39.8%; Score 43; DB 1; Length 2512;
Best Local Similarity 55.6%; Pred. No. 1.7e+02;
Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 3 NPNDYLISKDYGTATKVK 20
ID 1 1:|||||1
AC P27283;
DR 01-AUG-1992 (rel. 23, Created)

```

```

DT 01-AUG-1992 (rel. 23, last sequence update)
DT 16-OCT-2001 (rel. 40, last annotation update)
DE Nonstructural polyprotein (p270) [Contains: Nonstructural protein
DE NSP1; Nonstructural NSP2; Nonstructural NSP3; Nonstructural NSP4].
OS Sindbis virus (subtype Ockelbo / strain Edsbyn 82-5).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
OC Alphavirus.
OX NCBI_TaxID-31699;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-91220725; PubMed-1673813;
RA Shirako Y., Niklasson B., Dairymple J.M., Strauss E.G., Strauss J.H.;
RT "Structure of the Ockelbo virus genome and its relationship to other
  Sindbis viruses.";
RL Virology 182:753-764(1991).
CC -1- FUNCTION: NSP2 MAY BE INVOLVED IN RNA BINDING DURING REPLICATION.
CC -1- PFM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- MISCELLANEOUS: READTHROUGH OF THE TERMINATOR CODON UGA OCCURS
  BETWEEN THE CODONS FOR 1896-TYR AND 1897-LEU.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M69205; AAA6972.1; ALT_SEQ.
CC PIR: A39991; MNW82.
DR DR
DR MEROPS: C09.001;
DR InterPro: IPR002589; Alpp.
DR InterPro: IPR002620; Peptidase_C9.
DR InterPro: IPR001788; RNA_dep_RNAPol2.
DR InterPro: IPR002588; V_methyltransf.
DR InterPro: IPR000606; Viral_helicase1.
DR pfam: PF00978; RNA_dep_RNAPol2; 1.
DR pfam: PF01443; Viral_helicase1; 1.
DR pfam: PF01660; Vmethyltransf; 1.
DR pfam: PF01661; Alpp; 1.
DR pfam: PF01707; Peptidase_C9; 1.
DR SMART: SM00506; Alpp; 1.
KW Polyprotein; Nonstructural protein; RNA-binding; Helicase.
FT CHAIN 1 540 NONSTRUCTURAL PROTEIN NSP1.
FT CHAIN 541 1347 NONSTRUCTURAL PROTEIN NSP2.
FT CHAIN 1348 1896 NONSTRUCTURAL PROTEIN NSP3.
FT CHAIN 1897 2514 NONSTRUCTURAL PROTEIN NSP4.
SO SEQUENCE 2514 AA; 279642 MW; 2F388CF32ACF5EDD CRC64;

Query Match 39.8%; Score 43; DB 1; Length 2514;
Best Local Similarity 55.6%; Pred. No. 1.7e+02;
Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 3 NPNDYLISKDYGTATKVK 20
ID 1 1:|||||1
AC 877 NPCKKNIEIDITGATKPK 894
DR 01-AUG-1992 (rel. 23, Created)

```

Search completed: June 26, 2003, 00:50:34
Job time : 3.22353 secs

Thu Jun 26 11:58:22 2003

us-09-980-370-8.rsp

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 26, 2003, 00:47:39 ; Search time 6.42353 Seconds
(without alignments)

673.616 Million cell updates/sec

Title: US-09-980-370-8

Sequence: 1 ALPNPNYLSKDVGTATKVKY 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: SP-archaeap:
2: SP-bacteri:
3: SP-fungi:
4: SP-human:
5: SP-invertebrate:
6: SP-mammal:
7: SP-plant:
8: SP-protist:
9: SP-phage:
10: SP-plant:
11: SP-todent:
12: SP-virus:
13: SP-unclassified:
14: SP-unclassified:
15: SP-virus:
16: SP-bacteri:
17: SP-archaeap:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of this result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	Score	Query Length	ID	Description
1	108	100.0	1097	P2196 porphyromon
2	108	100.0	1223	P2196 porphyromon
3	108	100.0	1358	P66967 porphyromon
4	108	100.0	1587	P66967 porphyromon
5	108	100.0	1587	P66967 porphyromon
6	108	100.0	1706	P2196 porphyromon
7	108	100.0	1706	P2196 porphyromon
8	108	100.0	1723	P2196 porphyromon
9	108	100.0	1723	P2196 porphyromon
10	108	100.0	1723	P2196 porphyromon
11	108	100.0	1723	P2196 porphyromon
12	108	100.0	1723	P2196 porphyromon
13	64	59.3	925	O9F430 porphyromon
14	57	52.8	654	O9F430 porphyromon
15	55	50.9	334	O45985 neisseria
16	50	46.3	559	O9K058 neisseria

ALIGNMENTS

17	50	46.3	580	2	O9JPU5	O9JPU5 neisseria m
18	50	46.3	580	16	O9JPU22	O9JPU22 neisseria m
19	50	46.3	580	16	O9JPU21	O9JPU21 neisseria m
20	49	45.4	248	16	O9JPU21	O9JPU21 neisseria m
21	48	44.4	372	3	O9JPU21	O9JPU21 neisseria m
22	48	44.4	372	3	O9JPU21	O9JPU21 neisseria m
23	48	44.4	372	3	O9JPU21	O9JPU21 neisseria m
24	47.5	44.0	513	10	O9JPU21	O9JPU21 neisseria m
25	47	43.5	182	2	O9JPU21	O9JPU21 neisseria m
26	47	43.5	182	2	O9JPU21	O9JPU21 neisseria m
27	47	43.5	182	2	O9JPU21	O9JPU21 neisseria m
28	47	43.5	182	2	O9JPU21	O9JPU21 neisseria m
29	47	43.5	182	2	O9JPU21	O9JPU21 neisseria m
30	46	42.6	326	12	O9JPU21	O9JPU21 neisseria m
31	46	42.6	326	12	O9JPU21	O9JPU21 neisseria m
32	46	42.6	326	12	O9JPU21	O9JPU21 neisseria m
33	46	42.6	326	12	O9JPU21	O9JPU21 neisseria m
34	46	42.6	326	12	O9JPU21	O9JPU21 neisseria m
35	45	41.7	979	5	O9JPU21	O9JPU21 neisseria m
36	45	41.7	979	5	O9JPU21	O9JPU21 neisseria m
37	45	41.7	979	5	O9JPU21	O9JPU21 neisseria m
38	44.5	41.2	324	2	O9JPU21	O9JPU21 neisseria m
39	44.5	41.2	324	2	O9JPU21	O9JPU21 neisseria m
40	44.5	41.2	324	2	O9JPU21	O9JPU21 neisseria m
41	44.5	41.2	324	2	O9JPU21	O9JPU21 neisseria m
42	44	40.7	1080	16	O9JPU21	O9JPU21 neisseria m
43	44	40.7	315	10	O9JPU21	O9JPU21 neisseria m
44	44	40.7	320	16	O9JPU21	O9JPU21 neisseria m
45	44	40.7	390	16	O9JPU21	O9JPU21 neisseria m

RESULT 1
P2196 PRELIMINARY: PRT: 1097 AA.

AC P2196: 01-FEB-1997 (TREMBL) 02, Created
DT 01-FEB-1997 (TREMBL) 02, Last sequence update
DT 01-JUN-2002 (TREMBL) 21, Last annotation update
DE TONR-linked adhesin precursor.
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteri: CPB group: Bacteroidales; Bacteroidetes; Porphyromonadaceae;
OX NCBI_TaxID=837;
RN L1: JOURNAL: FROM N.A.
RX MEDLINE=91366416; PubMed=9244285;
RA Aduse-Opoku J., Stanley J.M., Young K.A., Mull J., Rangarajan M.,
Curtis M.A.;
RT "The tla gene of Porphyromonas gingivalis W50: a homologue of the
arginine-specific protease precursor (trpR) which shares sequence
homology to the tla gene of Bacteroides fragilis."
RT J Bacteriol 179:4778-4786(1997).
DR EMBL: Y07618; CAA68897.1; .
DR InterPro: IPR000977; DNM_119ase.
DR InterPro: IPR001769; Peptidase_C25.
DR Pfam: PF01364; Peptidase_C25.1.
DR KEGG: P500637; DNM_119ase; UNKNOWN_1.
FT SIGNAL 1 53 POTENTIAL.
SO SEQUENCE 1097 AA: 118731 MW: 7388A378421F8B9 CRC64:

Query Match Similarity 100.0% Score 108 DB 2: Length 1097:
Matches 21: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
QY 1 ALPNPNYLSKDVGTATKVKY 21
DB 581 ALPNPNYLSKDVGTATKVKY 601

RESULT 2

092NB5 PRELIMINARY: PRT: 1223 AA.
 ID 092NB5
 AC 092NB5
 DT 01-MAY-1999 (TREMBlrel. 10, Created)
 DT 01-JUN-2002 (TREMBlrel. 10, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE 130K-HMD (Fragment)
 DE Hemagglutinin/protease
 OS Porphyromonas gingivalis (Bacteroides gingivalis)
 OC Bacteri; CF8 group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
 CC Porphyromonas
 CC NCBI_TaxID=837;
 RN 11
 RP SEQUENCE FROM N.A.
 RC STRAIN=381;
 RA MEDLINE-99143166; PubMed-9988746;
 RA Shiba Y., Hayakawa M., Takiguchi H., Shirota T., Abiko Y.:
 RT "Cloning, sequencing and characterization of the hemagglutinin-associated
 RT short repeat found in Porphyromonas gingivalis multiple gene
 RT products."
 RL J. Biol. Chem. 274:5012-5020(1999).
 RL EMBL: AB019363; SAA34341.1;
 DR Interpro: IPRO00977; DML11998.
 DR HSP: P5493; ICVR.
 DR Pfam: PF01364; Peptidase_C25.
 DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
 FT NON_TER
 SQ SEQUENCE 1223 AA: 131542 MW: 00225CD2B8F9E1B3 CRC64;

Query Match 100.0%; Score 108; DR 2; Length 1223;
 Best Local Similarity 100.0%; Pred No. 1 2e-08;
 Matches 21: Conservative 0; Mismatches 0; Indels 0; Caps 0;

OY 1 ALNDPNTLSKDVGTGATKRVY 21
 DB 707 ALNDPNTLSKDVGTGATKRVY 727

RESULT 3

096967 PRELIMINARY: PRT: 1358 AA.
 ID 096967
 AC 096967
 DT 01-MAY-1997 (TREMBlrel. 03, Created)
 DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Hemagglutinin
 DE Hemagglutinin
 OS Porphyromonas gingivalis (Bacteroides gingivalis)
 OC Bacteri; CF8 group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
 CC Porphyromonas
 CC NCBI_TaxID=837;
 RN 11
 RP SEQUENCE FROM N.A.
 RC STRAIN=381;
 RA Han N., Leplante G., Wojciechowski L., Progniske-Fox A.:
 RT "Cloning, sequencing and characterization of hagd, a member of the
 RT hupE multigene family in Porphyromonas gingivalis."
 RT Submitted (Aug-1996) to the EMBL/Genbank/DBJ databases.
 RL EMBL: AF015281; AA69399.1;
 DR Interpro: IPRO00977; DML11998.
 DR HSP: P5493; ICVR.
 DR Pfam: PF01364; Peptidase_C25.
 DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
 SQ SEQUENCE 1358 AA: 147102 MW: 47C6A0B2550DEB8 CRC64;

Query Match 100.0%; Score 108; DR 2; Length 1358;
 Best Local Similarity 100.0%; Pred No. 1 3e-08;
 Matches 21: Conservative 0; Mismatches 0; Indels 0; Caps 0;

OY 1 ALNDPNTLSKDVGTGATKRVY 21
 DB 850 ALNDPNTLSKDVGTGATKRVY 870

RESULT 4

096987 PRELIMINARY: PRT: 1687 AA.
 ID 096987
 AC 096987
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Hemagglutinin/protease
 DE HAGE
 OS Porphyromonas gingivalis (Bacteroides gingivalis)
 OC Bacteri; CF8 group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
 CC Porphyromonas
 CC NCBI_TaxID=837;
 RN 11
 RP SEQUENCE FROM N.A.
 RC STRAIN=381;
 RA Han N., Dong H., Progniske-Fox A.:
 RT "Cloning, sequencing and characterization of hagd from P. gingivalis 381."
 RT Submitted (Sep-1999) to the EMBL/Genbank/DBJ databases.
 RL EMBL: AF026946; AAD01810.1;
 DR HSP: P5493; ICVR.
 DR MEROPS: C25.001;
 DR Interpro: IPRO00977; DML11998.
 DR HSP: P5493; ICVR.
 DR Pfam: PF01364; Peptidase_C25.
 DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
 KW Protease.
 SQ SEQUENCE 1687 AA: 183702 MW: D085916A39FE70 CRC64;

Query Match 100.0%; Score 108; DR 2; Length 1687;
 Best Local Similarity 100.0%; Pred No. 1 7e-08;
 Matches 21: Conservative 0; Mismatches 0; Indels 0; Caps 0;

OY 1 ALNDPNTLSKDVGTGATKRVY 21
 DB 1179 ALNDPNTLSKDVGTGATKRVY 1199

RESULT 5

051816 PRELIMINARY: PRT: 1704 AA.
 ID 051816
 AC 051816
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Arg-gingipain-1 proteinase
 OS Porphyromonas gingivalis (Bacteroides gingivalis)
 OC Bacteri; CF8 group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
 CC Porphyromonas
 CC NCBI_TaxID=837;
 RN 11
 RP SEQUENCE FROM N.A.
 RC STRAIN=5138080; PubMed-7836351;
 RA Travis J., Barr P.J., Pike R.N., Prochazka V., Klefer M.C.,
 RT "Molecular cloning and structural characterization of the Arg-
 RT gingipain proteinase of Porphyromonas gingivalis. Biosynthesis as a
 RT protease-inhibitor polypeptide."
 RL EMBL: AF015281; AA69399.1;
 DR HSP: P5493; ICVR.
 DR MEROPS: C25.001;
 DR Interpro: IPRO00977; DML11998.
 DR HSP: P5493; ICVR.
 DR Pfam: PF01364; Peptidase_C25.
 DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
 FT CHAIN 228
 FT MATURE 50-KDA CYSTEINE PROTEINASE
 FT CINGIPAIN
 SQ SEQUENCE 1704 AA: 185436 MW: 6A34A013CA2676 CRC64;

Query Match 100.0%; Score 108; DR 2; Length 1704;

Best Local Similarity 100.0%; Pred. No. 1.7e-08;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALNDPNYLISKDYTGATKVKY 21
DB 1196 ALNDPNYLISKDYTGATKVKY 1216

RESULT 6

OY1839 PRELIMINARY; PRT; 1706 AA.

AC O51839; O51840;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Arginine-specific thiol protease precursor.
CN PPR1.
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; CFb group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
OC Porphyromonas.
OX NCBI_TaxID=837;
RN SEQUENCE FROM N.A.
RC STRAIN-W50;
RX MEDLINE=95160709; PubMed=7857299;
RA Kirszenbaum L., Sotiropoulos C., Jackson C., Cleal S., Slakeski N., Reynolds E.C.;
RT "Complete nucleotide sequence of a gene pthr of Porphyromonas gingivalis W50 encoding a 132 kDa protein that contains an arginine-specific thiol endopeptidase domain and a haemagglutinin domain.";
RL Blochem. Biophys. Res. Commun. 207:424-431(1995).
RN [12]
RN SEQUENCE FROM N.A.
RC STRAIN-W50;
RX MEDLINE=96311339; PubMed=8713096;
RA Slakeski N., Cleal S.M., Reynolds E.C.;
RT "Characterization of a Porphyromonas gingivalis gene pthr that encodes an arginine-specific thiol proteinase and multiple adhesins.";
RL Blochem. Biophys. Res. Commun. 224:605-610(1996).
RN [13]
RN SEQUENCE FROM N.A.
RC STRAIN-W50;
RA Reynolds E.;
RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.
RN [14]
RN SEQUENCE FROM N.A.
RC STRAIN-W50;
RA Slakeski N.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL: L26341; AAC18876.1; -.
DR HSSP: P95493; 1CVR.
DR InterPro: IPR000977; DNA_ligase.
DR InterPro: IPR001769; Peptidase_C25.
DR Pfam: PF01364; Peptidase_C25.3.
DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
KM Protease; Signal.
FT SIGNAL 23
FT CHAIN 228 1706 ARGININE-SPECIFIC THIOLE PROTEASE.
SQ SEQUENCE 1706 AA; 185626 MW; E8BDF07C9813B844 CRC64;

Query Match 100.0%; Score 108; DB 2; Length 1706;
Best Local Similarity 100.0%; Pred. No. 1.7e-08;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALNDPNYLISKDYTGATKVKY 21
DB 1198 ALNDPNYLISKDYTGATKVKY 1218

RESULT 7
OY1838 PRELIMINARY; PRT; 1706 AA.
AC O51838;
OY1838;

DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-MAY-1997 (TREMblrel. 03, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Arginine-specific thiol protease precursor.
CN PPR1.
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; CFb group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
OC Porphyromonas.
OX NCBI_TaxID=837;
RN SEQUENCE FROM N.A.
RC STRAIN-W50;
RX MEDLINE=96071894; PubMed=7591131;
RA Aduse-Opoku J., Milr J., Slaney J.M., Rangarajan M., Curtis M.A.;
RT "Characterization, genetic analysis, and expression of a protease antigen (PPR1) of Porphyromonas gingivalis W50.";
RL Infect. Immun. 63:4744-4754(1995).
RN [12]
RN SEQUENCE FROM N.A.
RC STRAIN-W50;
RA Rangarajan M., Aduse-Opoku J., Slaney J.M., Young K.A., Curtis M.A.;
RT "The ppr1 and the ppr2 arginine-specific protease genes of Porphyromonas gingivalis W50 produce five biochemically distinct enzymes.";
RL Mol. Microbiol. 23:0-0(1997).
DR EMBL: X82680; CA57997.1; -.
DR HSSP: P95493; 1CVR.
DR MEROPS: C25.001; -.
DR InterPro: IPR000977; DNA_ligase.
DR InterPro: IPR001769; Peptidase_C25.
DR Pfam: PF01364; Peptidase_C25.3.
DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
KM Signal.
FT SIGNAL 1 23
FT CHAIN 228 719 ALPHA-PROTEASE.
FT CHAIN 720 1262 BETA-ADHESIN.
SQ SEQUENCE 1706 AA; 185705 MW; 0E56DC87EDAB8CDD CRC64;

Query Match 100.0%; Score 108; DB 2; Length 1706;
Best Local Similarity 100.0%; Pred. No. 1.7e-08;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALNDPNYLISKDYTGATKVKY 21
DB 1198 ALNDPNYLISKDYTGATKVKY 1218

RESULT 8

P72194 PRELIMINARY; PRT; 1723 AA.

AC P72194;
DT 01-FEB-1997 (TREMblrel. 02, Created)
DT 01-FEB-1997 (TREMblrel. 02, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Lys-gingipain.
GN KGP.
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; CFb group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
OC Porphyromonas.
OX NCBI_TaxID=837;
RN SEQUENCE FROM N.A.
RC STRAIN-W50;
RX MEDLINE=97044756; PubMed=8889827;
RA Okamoto K., Kadowaki T., Nakayama K., Yamamoto K.;
RT "Cloning and sequencing of the gene encoding a novel lysine-specific cysteine proteinase (Lys-gingipain) in Porphyromonas gingivalis: structural relationship with the arginine-specific cysteine proteinase (Arg-gingipain).";
RL J. Biochem. 120:398-406(1996).
DR EMBL: D83258; BAA11870.1; -.
DR MEROPS: C25.002; -.
DR InterPro: IPR000977; DNA_ligase.

DR InterPro: IPR001769; Peptidase_C25.
 DR Pfam: PF01364; Peptidase_C25; 3. UNKNOWN_1
 DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1
 SO SEQUENCE 1723 AA; 187261 MW; 5628963D231493EB CRC64;

Query Match 100.0%; Score 108; DB 2; Length 1723;
 Best Local Similarity 100.0%; Pred. No. 1.7e-08;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1 ALNPDNYLSKDYTGATKVKY 21
 1215 ALNPDNYLSKDYTGATKVKY 1235

RESULT 9
 ID P72197 PRELIMINARY; PRT; 1723 AA.
 AC P72197;
 DT 01-FEB-1997 (TREMBLrel. 02, Created)
 DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE lvs-gingipain.
 GN KCP.
 OS Porphyromonas gingivalis (Bacteroides gingivalis).
 OC Bacteria; CFB group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
 CC Porphyromonas.
 NCBI_TaxID=837;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Pavloff N., Pemberton P.A., Potempa J., Chen W.-C.A., Pike R.N.,
 RA Prochazka V., Kiefer M.C., Travis J., Barr P.J.;
 RT "Molecular cloning and characterization of Porphyromonas gingivalis
 RT lvs-gingipain.";
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL:U04691; AAA9810.1; -.
 DR MEROPS: C25.002; -.
 DR InterPro: IPR000977; DNA_Ligase.
 DR Pfam: PF01364; Peptidase_C25; 3.
 DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
 SO SEQUENCE 1723 AA; 186831 MW; 4508A7E50197CEBD CRC64;

Query Match 100.0%; Score 108; DB 2; Length 1723;
 Best Local Similarity 100.0%; Pred. No. 1.7e-08;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALNPDNYLSKDYTGATKVKY 21
 1215 ALNPDNYLSKDYTGATKVKY 1235

RESULT 10
 ID 052050 PRELIMINARY; PRT; 1732 AA.
 AC 052050;
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Lysine specific cysteine protease.
 OS Porphyromonas gingivalis (Bacteroides gingivalis).
 OC Bacteria; CFB group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
 CC Porphyromonas.
 NCBI_TaxID=837;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=883;
 RX MEDLINE=98298016; PubMed=9632563;
 RA Lewis J.P., Macrina F.L.;
 RT "IS195, an insertion sequence-like element associated with protease
 RT genes in Porphyromonas gingivalis.";
 RL Infect. Immun. 66:3035-3042(1998).
 DR EMBL: AF017059; AAC26523.1; -.
 DR MEROPS: C25.002; -.

DR InterPro: IPR000977; DNA_Ligase.
 DR InterPro: IPR001769; Peptidase_C25.
 DR Pfam: PF01364; Peptidase_C25; 3. UNKNOWN_1
 DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1
 SO SEQUENCE 1732 AA; 187931 MW; B2337463D5CB5EAS CRC64;

Query Match 100.0%; Score 108; DB 2; Length 1732;
 Best Local Similarity 100.0%; Pred. No. 1.7e-08;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1 ALNPDNYLSKDYTGATKVKY 21
 1216 ALNPDNYLSKDYTGATKVKY 1236

RESULT 11
 ID 051817 PRELIMINARY; PRT; 1732 AA.
 AC 051817;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Porphyryain.
 GN PRP.
 OS Porphyromonas gingivalis (Bacteroides gingivalis).
 OC Bacteria; CFB group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
 CC Porphyromonas.
 NCBI_TaxID=837;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=M12;
 RX MEDLINE=96213011; PubMed=8631659;
 RA Barkocy-Gallagher G.A., Han N., Patti J.M., Whitlock J.,
 RA Prokuse-fox A., Lantz M.S.;
 RT "Analysis of the prt gene encoding porphyryain, a cysteine proteinase
 RT of Porphyromonas gingivalis.";
 RL J. Bacteriol. 178:2734-2741(1996).
 DR EMBL: U42210; AAB06565.1; -.
 DR MEROPS: C25.002; -.
 DR InterPro: IPR000977; DNA_Ligase.
 DR Pfam: PF01364; Peptidase_C25; 3.
 DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_1.
 SO SEQUENCE 1732 AA; 187875 MW; 654271DBEF7BCA64 CRC64;

Query Match 100.0%; Score 108; DB 2; Length 1732;
 Best Local Similarity 100.0%; Pred. No. 1.7e-08;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALNPDNYLSKDYTGATKVKY 21
 1216 ALNPDNYLSKDYTGATKVKY 1236

RESULT 12
 ID 007442 PRELIMINARY; PRT; 1732 AA.
 AC 007442;
 DT 01-JUL-1997 (TREMBLrel. 04, Created)
 DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Lysine-specific cysteine proteinase.
 GN PRK.
 OS Porphyromonas gingivalis (Bacteroides gingivalis).
 OC Bacteria; CFB group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
 CC Porphyromonas.
 NCBI_TaxID=837;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=850;
 RX MEDLINE=99235907; PubMed=10219167;
 RA Stakeki N., Cleal S.M., Bhogal P.S., Reynolds E.C.;

RT *Characterization of a Porphyromonas gingivalis gene prtK that encodes
 RT a lysine-specific cysteine proteinase and three sequence-related
 RT adhesins."
 RL Oral Microbiol. Immunol. 14:92-97(1999).
 DR EMBL: U75366; AAB60809.1; -.
 DR MEROPS: C25.002; -.
 DR InterPro: IPR000977; DNA_Ligase.
 DR InterPro: IPR001769; Peptidase_C25.
 DR Pfam: PF01364; Peptidase_C25; 3.
 DR PROSITE: PS00697; DNA_Ligase_A1; UNKNOWN_1.
 SQ SEQUENCE 1732 AA; 187914 MW; 45D5B9137391703 CRC64;

Query Match 100.0%; Score 108; DB 2; Length 1732;
 Best Local Similarity 100.0%; Pred. No. 1.7e-08;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALNPDNYLISKDVGTGATKVKY 21
 DB 1216 ALNPDNYLISKDVGTGATKVKY 1236

RESULT 13
 OY 09F4J0 PRELIMINARY: PRT; 925 AA.

AC 09F4J0:
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Putative outer membrane protein PG57.
 OS Porphyromonas gingivalis (Bacteroides gingivalis).
 OC Bacteria; CFb group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
 OC Porphyromonas.
 NCBI_TaxID=837;
 RN NCBI
 RP SEQUENCE FROM N.A.
 RC STRAIN=W50;
 RA Ross B.C., Barr I., Patterson M., Agius C., Rothel L., Margelits M.,
 RT Hocking D., Webb E.;
 RT P. gingivalis polypeptides and nucleic acids."
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=W50;
 RA Ross B.C., Czajkowski L., Hocking D., Margelits M., Webb E., Rothel L.,
 RT Patterson M., Agius C., Camuglia S., Reynolds E., Barr I.G.;
 RT "Identification of vaccine candidates from genomic analysis of
 RT Porphyromonas gingivalis."
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY007522; AAG24228.1; -.
 DR InterPro: IPR003961; FN_III.
 DR SMART: SM00060; FN3; 3.
 SQ SEQUENCE 925 AA; 103632 MW; 5FF2198D6914DAE8 CRC64;

Query Match 59.3%; Score 64; DB 2; Length 925;
 Best Local Similarity 60.0%; Pred. No. 0.12;
 Matches 12; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 2 LNPDPNYLISKDVGTGATKVKY 21
 DB 367 VTPDPNYLITPKYEGAKRKVKY 386

RESULT 14
 OY 093922 PRELIMINARY: PRT; 654 AA.
 AC 093922:
 DT 01-MAY-1999 (TREMBlrel. 10, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE El-Like protein.
 GN GSA7.
 OS Pichia pastoris (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Saccharomycetaceae; Pichia.
 OX NCBI_TaxID=4922;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99250148; PubMed=10233149;
 RA Yuan W., Stromhaug P.E., Dunn W.A. Jr.;
 RT "Glucose-induced autophagy of peroxisomes in Pichia pastoris requires
 RT a unique El-Like protein."
 RL Mol. Biol. Cell 10:1353-1366(1999).
 DR EMBL: AF098976; AAD14610.1; -.
 DR InterPro: IPR000205; NMD_binding.
 DR InterPro: IPR000594; Thif_domain.
 DR Pfam: PF00899; Thif; 1.
 SQ SEQUENCE 654 AA; 74292 MW; 893937F5FB30D8F CRC64;

Query Match 52.8%; Score 57; DB 3; Length 654;
 Best Local Similarity 65.0%; Pred. No. 1.2;
 Matches 13; Conservative 3; Mismatches 2; Indels 2; Gaps 1;

OY 1 ALNPDNYLISKDVGTGATKVKY 20
 DB 609 ALNENNYL-EDLTGLTKVK 626

RESULT 15
 OY 045985 PRELIMINARY: PRT; 334 AA.

AC 045985:
 DT 01-JUN-1998 (TREMBlrel. 06, Created)
 DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE ZK1037.3 protein.
 GN ZK1037.3.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN NCBI
 RP SEQUENCE FROM N.A.
 RA Basham V.M.;
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode C.elegans: A platform for
 RT investigating biology."
 RL Science 282:2012-2018(1998).
 DR EMBL: Z81142; CAB03502.1; -.
 DR InterPro: IPR002651; DUF32.
 DR Pfam: PF01748; DUF32; 1.
 SQ SEQUENCE 334 AA; 37719 MW; E8592D5A4D24FBA3 CRC64;

Query Match 50.9%; Score 55; DB 5; Length 334;
 Best Local Similarity 64.7%; Pred. No. 1.2;
 Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 2 LNPDPNYLISKDVGTGATK 18
 DB 22 LNPDPVCSDDVTSHTK 38

Search completed: June 26, 2003, 00:53:27
 Job time : 7.42353 secs

XX The present invention relates to prophylaxis or treatment of infection
 CC by a microorganism in a biological environment comprising iron, heme
 CC or porphyrin. The treatment involves administering to the environment,
 CC an agent which antagonizes the interaction between a molecule derived
 CC from the microorganism and having an HA2 domain, and an HA2-binding
 CC motif on a porphyrin containing molecule present in the environment.
 CC Useful in the manufacture of a medicament for the prophylaxis and
 CC treatment of periodontal, pulmonary, vaginal, urethral or hoof disease
 CC resulting from Porphyromonas gingivalis infection or infection by a
 CC related microorganism.

XX Sequence 15 AA:

Query Match 100.0%; Score 78; DB 22; Length 15;
 Best Local Similarity 100.0%; Pred. No. 1.3e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALNDPNYLISKDVTG 15
 DB 1 ALNDPNYLISKDVTG 15

RESULT 2

ID AAB49218 standard; peptide: 21 AA.

AC AAB49218;

DT 13-MAR-2001 (first entry)

DE Peptide #3.

KW HA2 domain; porphyrin; periodontal; pulmonary; vaginal; urethral;
 infection.

OS Unidentified.

PN WO200072875-A1.

XX 07-DEC-2000.

PF 26-MAY-2000; 2000WO-AU00599.

PR 28-MAY-1999; 99AU-0000652.

XX (UNSY) UNITV SYDNEY.

PI Collyer CA, Hunter N, De Carlo AA;

DR WPI: 2001-080424/09.

XX Treating microbial infection in environment containing porphyrin, by
 PT administering a HA-2 antagonist -

PS Claim 11; Page 40; 102pp; English.

XX The present invention relates to prophylaxis or treatment of infection
 CC by a microorganism in a biological environment comprising iron, heme
 CC or porphyrin. The treatment involves administering to the environment,
 CC an agent which antagonizes the interaction between a molecule derived
 CC from the microorganism and having an HA2 domain, and an HA2-binding
 CC motif on a porphyrin containing molecule present in the environment.
 CC Useful in the manufacture of a medicament for the prophylaxis and
 CC treatment of periodontal, pulmonary, vaginal, urethral or hoof disease
 CC resulting from Porphyromonas gingivalis infection or infection by a
 CC related microorganism.

XX Sequence 21 AA:

Query Match 100.0%; Score 78; DB 22; Length 21;
 Best Local Similarity 100.0%; Pred. No. 2e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALNDPNYLISKDVTG 15
 DB 1 ALNDPNYLISKDVTG 15

RESULT 3

ID AAB49217 standard; protein: 134 AA.

AC AAB49217;

DT 13-MAR-2001 (first entry)

DE Peptide used in the invention.

KW HA2 domain; porphyrin; periodontal; pulmonary; vaginal; urethral;
 infection.

OS Unidentified.

PN WO200072875-A1.

XX 07-DEC-2000.

PF 26-MAY-2000; 2000WO-AU00599.

PR 28-MAY-1999; 99AU-0000652.

XX (UNSY) UNITV SYDNEY.

PI Collyer CA, Hunter N, De Carlo AA;

DR WPI: 2001-080424/09.

XX Treating microbial infection in environment containing porphyrin, by
 PT administering a HA-2 antagonist -

PS Claim 9; Page 98-99; 102pp; English.

XX The present invention relates to prophylaxis or treatment of infection
 CC by a microorganism in a biological environment comprising iron, heme
 CC or porphyrin. The treatment involves administering to the environment,
 CC an agent which antagonizes the interaction between a molecule derived
 CC from the microorganism and having an HA2 domain, and an HA2-binding
 CC motif on a porphyrin containing molecule present in the environment.
 CC Useful in the manufacture of a medicament for the prophylaxis and
 CC treatment of periodontal, pulmonary, vaginal, urethral or hoof disease
 CC resulting from Porphyromonas gingivalis infection or infection by a
 CC related microorganism.

XX Sequence 134 AA:

Query Match 100.0%; Score 78; DB 22; Length 134;
 Best Local Similarity 100.0%; Pred. No. 1.9e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALNDPNYLISKDVTG 15
 DB 60 ALNDPNYLISKDVTG 74

RESULT 4

ID AAR96024 standard; Protein: 439 AA.

XX AAR96024;

DT 04-SEP-1996 (first entry)

DE P. gingivalis haemagglutinin hAgA HAREP4 product.

XX Haemagglutinin; hAgA; periodontal disease; vaccine; antibody;

KM Harep4.
 OS Porphyromonas gingivalis strain 381.
 XX
 PN MO9617936-A2.
 XX
 PD 13-JUN-1996.
 XX
 PF 11-DEC-1995; 95WO-US16108.
 XX
 PR 09-DEC-1994; 94US-0353485.
 XX
 PA (UABR-) UAB RES FOUND.
 PA (UYFL) UNIV FLORIDA.
 XX
 PI Han N, Lantz M, Lepine G, Patti JM, Proguliske-Fox A;
 PI Tumwasorn S;
 PI Tumwasorn S;
 XX
 DR WPI; 1996-287181/29.
 DR N-PSDB; AAT30648.
 XX
 PT Porphyromonas gingivalis genes and proteins - used in the detection
 PT and vaccination against periodontal disease
 XX
 PS Claim 4; Page 114-115; 153pp; English.
 XX
 CC HAREP4 (AAR96024) is the product of the HAREP4 repeat unit (AAT30648)
 CC of the hga gene (AAT30654) of P. gingivalis 318. It forms part
 CC of haemagglutinin hga (see also AAR96030). HAREP4 and other
 CC hga repeat unit products (see also AAR96021-23) can be obtd. from
 CC transformed host cells and used as vaccines to protect humans or
 CC animals against periodontal disease. Expression in Salmonella
 CC cells allows produ. of live vaccine. HAREP4 can also be used
 CC to detect the presence of anti-P. gingivalis antibodies and to
 CC raise monoclonal antibodies for diagnostic appln.
 XX
 SQ Sequence 439 AA;
 Query Match 100.0%; Score 78; DB 17; Length 439;
 Best Local Similarity 100.0%; Pred. No. 8e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 ALNPDNYLSKDYTG 15
 DB 251 ALNPDNYLSKDYTG 265
 RESULT 5
 ID AAM69492 standard; Protein: 439 AA.
 XX
 AC AAM69492;
 XX
 DT 22-DEC-1998 (first entry)
 XX
 DE Haemagglutinin protein hga, HAREP4.
 XX
 KM Haemagglutinin protein; periodontal disease; vaccine; hga.
 KM
 OS Porphyromonas gingivalis.
 OS
 XX US5824791-A.
 XX
 PN 20-OCT-1998.
 XX
 PD 11-DEC-1995; 95US-0570311.
 XX
 PF 11-DEC-1995; 95US-0570311.
 XX
 PR 08-SEP-1988; 88US-0241640.
 PR 25-JAN-1991; 91US-0647119.
 PR 09-DEC-1994; 94US-0353485.
 XX
 PA (UABR-) UAB RES FOUND.

PA (UYFL) UNIV FLORIDA.
 XX
 PI Han N, Lantz M, Lepine G, Patti JM, Proguliske-Fox A;
 PI Tumwasorn S;
 PI Tumwasorn S;
 XX
 DR WPI; 1998-582627/49.
 DR N-PSDB; AAV58879.
 XX
 PT Isolated Porphyromonas gingivalis genes - encoding haemagglutinin
 PT and/or protease poly(peptide(s))
 XX
 PS Claim 1; Column 139-144; 101pp; English.
 XX
 CC This sequence is encoded by a Porphyromonas gingivalis gene of the
 CC invention. This sequence represents the hga haemagglutinin protein. The
 CC polypeptides are used to produce antibodies to organisms associated with
 CC periodontal disease. The antibodies are also used in purification and
 CC identification procedures. The genes and polypeptides are used as
 CC vaccines against periodontal disease.
 XX
 SQ Sequence 439 AA;
 Query Match 100.0%; Score 78; DB 19; Length 439;
 Best Local Similarity 100.0%; Pred. No. 8e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 ALNPDNYLSKDYTG 15
 DB 251 ALNPDNYLSKDYTG 265
 RESULT 6
 ID AAR96021 standard; Protein: 450 AA.
 XX
 AC AAR96021;
 XX
 DT 04-SEP-1996 (first entry)
 XX
 DE P. gingivalis haemagglutinin hga HAREP1 product.
 XX
 KM Haemagglutinin; hga; periodontal disease; vaccine; antibody;
 KM HAREP1.
 XX
 OS Porphyromonas gingivalis strain 381.
 OS
 XX MO9617936-A2.
 XX
 PN 13-JUN-1996.
 PD 11-DEC-1995; 95WO-US16108.
 XX
 PF 11-DEC-1995; 95WO-US16108.
 XX
 PR 09-DEC-1994; 94US-0353485.
 XX
 PA (UABR-) UAB RES FOUND.
 PA (UYFL) UNIV FLORIDA.
 XX
 PI Han N, Lantz M, Lepine G, Patti JM, Proguliske-Fox A;
 PI Tumwasorn S;
 PI Tumwasorn S;
 XX
 DR WPI; 1996-287181/29.
 DR N-PSDB; AAT30645.
 XX
 PT Porphyromonas gingivalis genes and proteins - used in the detection
 PT and vaccination against periodontal disease
 XX
 PS Claim 4; Page 103-104; 153pp; English.
 XX
 CC HAREP1 (AAR96021) is the product of the HAREP1 repeat unit (AAT30645)
 CC of the hga gene (AAT30654) of P. gingivalis 318. It forms part
 CC of haemagglutinin hga (see also AAR96030). HAREP1 and other
 CC hga repeat unit products (see also AAR96022-24) can be obtd. from
 CC transformed host cells and used as vaccines to protect humans or

CC	Incorporating AG-1.	XX
SO	Sequence	1704 AA:
Alignment Scores:		
Pred. No.:	2,92e-77	Length: 1704
Score:	716.00	Matches: 131
Percent Similarity:	99.25%	Conservative: 2
Best Local Similarity:	97.76%	Mismatches: 1
Query Match:	97.95%	Indels: 0
DB:	16	Gaps: 0
US-09-980-370-5 (1-402) x AAR70188 (1-1704)		
OY	1 GCAGACTTCAGGAAAGCTTCGAGCTCTTACATCATGAGAGGACGACCGAGTGGACT	60
Db	1137 AIAAphehrhglutnrhphgilserserthrhisgluAlaProAlaGluTrpThr	1156
OY	61 ACTATCGATGCCGATGGCGATGTGAGAGGTGGCTCTGTCTGTCTCCGACAATTGGAC	120
Db	1157 ThrileasphlaaspglyaspGlyGlnGlyTrpLeucysleuSerSerGlyGlnLeuasp	1176
OY	121 TGGCTTACAGCTCTATGGCGGACCAAGCTACTAGCTCTTCTCATGAAATGGCT	180
Db	1177 TrpLeuThrAlaHisGlyGlyTyrHisAsnValAlaIAserPheSerTrpAsnGlyMetAla	1196
OY	181 TTGAAATCCGTAACTATCTCATCTCTCAAGGATTTACAGGCGCAGACGAAGTAAAGTAC	240
Db	1197 LeuAsnProAspAsnTyrLeuIleSerLysaspValThrGlyIleThrLysValLysTyr	1216
OY	241 TACTATTCAGTCAACAGACGCTTTCCGGGGATCACTATGCGGTGATGATCTCCAAGACG	300
Db	1217 TyrTrpAlaValAlaAsnAspGlyPheProGlyAspHisTyrAlaValMetIleSerLysThr	1236
OY	301 GGCACGACGCGCGGAGACTTCACGGTGTCTTTCGACGAAGCCCTTAACGAATAAATPAG	360
Db	1237 GlyHisAsnAlaGlyAspPheThrValValPheGluGluThrProAsnGlyIleAsnLys	1256
OY	361 GCGGAGACAGATTGCGTCTTTCCACGGAACCCAAATGGCGCC	402
Db	1257 GlyGlyAlaAspArgPheGlyLeuSerThrGluAlaAsnGlyAla	1270
RESULT 5		
AAW34843		
ID	AAW34843 standard; protein: 1704 AA.	
XX	AAW34843:	
AC		
XX	03-JUN-1998 (first entry)	
DT		
XX	Arg-gingipain high molecular weight prepolypeptide sequence.	
DE		
XX	Arg-specific gingipain protease; gingivalis: periodontal disease;	
KW	vaccine; infection.	
RW		
XX	Porphyromonas gingivalis.	
OS		
XX		
PH	Key	Location/Qualifiers
FT	Protein	1..227
XX		/note="precursor protein"
XX	MO9734629-31.	
PD	25-SEP-1992.	
XX		
PF	-21-MAR-1997:	97WO-US04635.
XX		
PR	22-MAR-1996:	96US-0013945.
XX		
PA	(MORE-) MOREHOUSE SCHOOL MEDICINE.	
PA	(UYGE-) UNIV GEORGIA RES FOUND INC.	
XX		
PI	Genco CA, Potempa J, Travis J, Genco C;	

XX WPI: 1997-479993/44.
 DR N-PSDB: AAT93872.
 XX Porphyromonas gingivalis Arg-specific gingipain protease peptide(s)
 PT - useful for protecting animals and humans from gingivitis and
 PT periodontal diseases
 XX
 PS Disclosure: Pages 68-73; 95pp: English.
 XX
 CC The present sequence represents an arginine-specific protease of
 CC Porphyromonas gingivalis. The following peptides, derived from Arg-
 CC and lys-specific high molecular weight proteases, offer protection
 CC against infection: YTYVYRDGK IKKGLATTE DGVATGNNE YCVRKYTAGS VSPVVC
 CC (I); YTPVEEKNG RMVIVAKRY (II); QLEPFDVAC VNGDFEEMP CEAFMLNRQ
 CC (III); GEPNPQAPVS NLATATGQOK VTLKMDAPSTR (IV); GNHCEVEYK YTAGVSPKVC
 CC KDVTY (V); RMENYEPGR YTPVEEKONG (VI); TEAGEEDYK RMFMYNEPGR (VII);
 CC DYTYVYRGTG KIKIGGLATTE TFEEDGVATG NMEYCVCAKY TAGVSPKVC (VIII); YTYVYRDGT
 CC KIKIGGLATTE EEDG (IX); RDGTRIKEIK TATFEEDGV ATGN (X); KIKIGLATTT
 CC FEEDGVATGN HEY (XI); KMDANGTPN PNPN PNPN POTTLSL (XII); and
 CC YTPVEEKNG RMVIVAKRY (XIII). They are used in vaccines to protect
 CC animals, including humans, from gingivitis and/or periodontal
 CC diseases.
 CC
 XX
 SQ Sequence 1704 AA:
 Alignment Scores:
 Pred. No.: 2,92e-77 Length: 1704
 Score: 716.00 Matches: 131
 Percent Similarity: 99.25% Conserved: 2
 Best Local Similarity: 97.76% Mismatches: 1
 Query Match: 97.95% Indels: 0
 Gaps: 0
 US-09-980-370-5 (1-402) x AAW34843 (1-1704)
 QY 1 GCACACTTCACGGGAAACGTTTCGAGTCTTCTACTATGAGNAGGACACCGCGGATGCAC 60
 Db 1137 AAlasphethrGluThrPheGlnSerSerThrHisGluAlaProAlaGluThr 1156
 QY 61 ACTATCGATGCGGATGCGAGATGAGGAGGTTGCGCTGTCTGTCTCCGACAAATTGCAC 120
 Db 1157 ThrIleaspsAlaaspGlyaspGlyGlnGlyThrPleucysLeuSerSerGlyGlnLeuasp 1176
 QY 121 TGCGTCAACAGCTCATGGCGGACCAACAGCTAGTAAGCTCTTTCATAGGAATGAGTGCCT 180
 Db 1177 TrpLeuThrAlaHisGlyGlyTThrAsnValAlaIleSerPheSerTrpAsnGlyMetAla 1196
 QY 181 TTGAATCGTGAATCACTATCTCATGTCMAAGATGTATACGGCGGACCAAGATGAAGTAC 240
 Db 1197 LeuasnProAspAsnTyrLeuIleSerLysaspValThrGlyAlaThrLysValLysTyr 1216
 QY 241 TACTATCCAGTCAACGACGAGGTTTCCCGGGGATCACTATGCGGTGATGATCTCCAAAGAC 300
 Db 1217 TyrTyrAlaValAlaAsnaspGlyPheProGlyaspHisTyrAlaValMetIleSerLysThr 1236
 QY 301 GGCACGAAGCGCGGAGACTTTCACGCTGTTTTCGAAAGAAAGCCCTACAGGAATAATTAAG 360
 Db 1237 GlyThrAsnAlaGlyAspPheThrValValAlaPheGlnGluGluThrProAsnGlyIleAsnLys 1256
 QY 361 GGCAGAGCAAGATTCGGTCTTTCACGAGCAAGCAAGCAATGCGCGC 402
 Db 1257 GlyGlyAlaIleArgPheGlyLeuSerThrGlnAlaAsnGlyAla 1270
 RESULT 6
 AAY67396
 ID AAY67396 standard; Protein: 1704 AA.
 AC AAY67396;
 XX 25-APR-2000 (first entry)
 DT Arg-gingipain-2 amino acid sequence.
 DE

XX The present invention relates to prophylaxis or treatment of infection
 CC by a microorganism in a biological environment comprising iron, heme
 CC or porphyrin. The treatment involves administering to the environment,
 CC an agent which antagonizes the interaction between a molecule derived
 CC from the microorganism and having an HA2 domain, and an HA2-binding
 CC motif on a porphyrin containing molecule present in the environment.
 CC Useful in the manufacture of a medicament for the prophylaxis and
 CC treatment of periodontal, pulmonary, vaginal, urethral or hoof disease
 CC resulting from Porphyromonas gingivalis infection or infection by a
 CC related microorganism.
 XX SQ Sequence 15 AA;

Query Match 100.0%; Score 78; DB 22; Length 15;
 Best Local Similarity 100.0%; Pred. No. 1.3e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNPDNYLISKDVTG 15
 DB 1 ALNPDNYLISKDVTG 15

RESULT 2
 AAB49218
 ID AAB49218 standard; peptide: 21 AA.

XX AC AAB49218;
 XX DT 13-MAR-2001 (first entry)

XX PEptide #3.

XX HA2 domain; porphyrin; periodontal; pulmonary; vaginal; urethral;
 XX infection.

XX Unidentified.

XX WO200072875-A1.

XX 07-DEC-2000.

XX 26-MAY-2000; 2000WO-AU00599.

XX 28-MAY-1999; 99AU-0000652.

XX (UNSY) UNIV SYDNEY.

XX Collyer CA, Hunter N, De Carlo AA;

XX WPI; 2001-080424/09.

XX Treating microbial infection in environment containing porphyrin, by
 XX administering a HA-2 antagonist -
 XX Claim 11; Page 40; 102pp; English.

XX The present invention relates to prophylaxis or treatment of infection
 CC by a microorganism in a biological environment comprising iron, heme
 CC or porphyrin. The treatment involves administering to the environment,
 CC an agent which antagonizes the interaction between a molecule derived
 CC from the microorganism and having an HA2 domain, and an HA2-binding
 CC motif on a porphyrin containing molecule present in the environment.
 CC Useful in the manufacture of a medicament for the prophylaxis and
 CC treatment of periodontal, pulmonary, vaginal, urethral or hoof disease
 CC resulting from Porphyromonas gingivalis infection or infection by a
 CC related microorganism.

XX SQ Sequence 21 AA;

Query Match 100.0%; Score 78; DB 22; Length 21;
 Best Local Similarity 100.0%; Pred. No. 2e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNPDNYLISKDVTG 15
 DB 1 ALNPDNYLISKDVTG 15
 RESULT 3
 AAB49217
 ID AAB49217 standard; protein; 134 AA.
 XX AC AAB49217;

XX DT 13-MAR-2001 (first entry)

XX PEptide used in the invention.

XX HA2 domain; porphyrin; periodontal; pulmonary; vaginal; urethral;
 XX infection.

XX Unidentified.

XX WO200072875-A1.

XX 07-DEC-2000.

XX 26-MAY-2000; 2000WO-AU00599.

XX 28-MAY-1999; 99AU-0000652.

XX (UNSY) UNIV SYDNEY.

XX Collyer CA, Hunter N, De Carlo AA;

XX WPI; 2001-080424/09.

XX Treating microbial infection in environment containing porphyrin, by
 XX administering a HA-2 antagonist -

XX Claim 9; Page 98-99; 102pp; English.

XX The present invention relates to prophylaxis or treatment of infection
 CC by a microorganism in a biological environment comprising iron, heme
 CC or porphyrin. The treatment involves administering to the environment,
 CC an agent which antagonizes the interaction between a molecule derived
 CC from the microorganism and having an HA2 domain, and an HA2-binding
 CC motif on a porphyrin containing molecule present in the environment.
 CC Useful in the manufacture of a medicament for the prophylaxis and
 CC treatment of periodontal, pulmonary, vaginal, urethral or hoof disease
 CC resulting from Porphyromonas gingivalis infection or infection by a
 CC related microorganism.

XX SQ Sequence 134 AA;

Query Match 100.0%; Score 78; DB 22; Length 134;
 Best Local Similarity 100.0%; Pred. No. 1.9e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNPDNYLISKDVTG 15
 DB 60 ALNPDNYLISKDVTG 74

RESULT 4
 AAR96024
 ID AAR96024 standard; Protein; 439 AA.
 XX AC AAR96024;

XX DT 04-SEP-1996 (first entry)

XX P. gingivalis haemagglutinin hagA HAREP4 product.
 XX Haemagglutinin; hagA; periodontal disease; vaccine; antibody;

KW Harep4.
 XX Porphyromonas gingivalis strain 381.
 OS
 XX WO9617936-A2.
 PN
 XX 13-JUN-1996.
 PD
 XX 11-DEC-1995; 95WO-US16108.
 PF
 XX 09-DEC-1994; 94US-0353485.
 PR
 XX (UABR-) UAB RES FOUND.
 PA (UYFL) UNIV FLORIDA.
 XX
 XX Han N, Lantz M, Lepine G, Patti JM, Progulskie-Fox A;
 PI Tumwasorn S;
 XX WPI; 1996-287181/29.
 DR N-PSDB; AAT30648.
 DR
 XX Porphyromonas gingivalis genes and proteins - used in the detection
 PT and vaccination against periodontal disease
 PS Claim 4; Page 114-115; 153pp; English.
 XX
 XX Harep4 (AAR96024) is the product of the Harep4 repeat unit (AAT30648)
 CC of the hga gene (AAT30654) of P. gingivalis 318. It forms part
 CC of haemagglutinin hgaA (see also AAR96030). Harep4 and other
 CC hgaA repeat unit products (see also AAR96021-23) can be obtd. from
 CC transformed host cells and used as vaccines to protect humans or
 CC animals against periodontal disease. Expression in Salmonella
 CC cells allows prodn. of live vaccine. Harep1-4 can also be used
 CC to detect the presence of anti-P. gingivalis antibodies and to
 CC raise monoclonal antibodies for diagnostic appln.
 XX
 XX Sequence 439 AA;
 SQ
 Query Match 100.0%; Score 78; DB 17; Length 439;
 Best Local Similarity 100.0%; Pred. No. 8e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ALNPDNYLISKDVTG 15
 DB 251 ALNPDNYLISKDVTG 265
 |||||
 RESULT 5
 AAW69492
 ID AAW69492 standard; Protein; 439 AA.
 AC
 XX AAW69492;
 XX
 DT 22-DEC-1998 (first entry)
 DE Haemagglutinin protein hgaA, Harep4.
 XX
 KW Haemagglutinin protein; periodontal disease; vaccine; hgaA.
 XX
 OS Porphyromonas gingivalis.
 XX
 PN US5824791-A.
 XX
 XX 20-OCT-1998.
 PD
 XX 11-DEC-1995; 95US-0570311.
 PF
 XX 11-DEC-1995; 95US-0570311.
 PR 08-SEP-1988; 88US-0241640.
 PR 25-JAN-1991; 91US-0647119.
 PR 09-DEC-1994; 94US-0353485.
 XX
 XX (UABR-) UAB RES FOUND.
 PA

PA (UYFL) UNIV FLORIDA.
 XX Han N, Lantz M, Lepine G, Patti JM, Progulskie-Fox A;
 PI Tumwasorn S;
 XX WPI; 1998-582627/49.
 DR N-PSDB; AAV58879.
 XX
 XX Isolated Porphyromonas gingivalis genes - encoding haemagglutinin
 PT and/or protease poly:peptide(s)
 XX Claim 1; Column 139-144; 101pp; English.
 PS
 XX This sequence is encoded by a Porphyromonas gingivalis gene of the
 CC invention. This sequence represents the hgaA haemagglutinin protein. The
 CC polypeptides are used to produce antibodies to organisms associated with
 CC periodontal disease. The antibodies are also used in purification and
 CC identification procedures. The genes and polypeptides are used as
 CC vaccines against periodontal disease.
 XX
 XX Sequence 439 AA;
 SQ
 Query Match 100.0%; Score 78; DB 19; Length 439;
 Best Local Similarity 100.0%; Pred. No. 8e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ALNPDNYLISKDVTG 15
 DB 251 ALNPDNYLISKDVTG 265
 |||||
 RESULT 6
 AAR96021
 ID AAR96021 standard; Protein; 450 AA.
 XX
 XX AAR96021;
 XX
 DT 04-SEP-1996 (first entry)
 DE P. gingivalis haemagglutinin hgaA Harep1 product.
 XX
 KW Haemagglutinin; hgaA; periodontal disease; vaccine; antibody;
 KW Harep1.
 XX
 OS Porphyromonas gingivalis strain 381.
 XX
 PN WO9617936-A2.
 XX
 PD 13-JUN-1996.
 PF 11-DEC-1995; 95WO-US16108.
 XX
 XX 09-DEC-1994; 94US-0353485.
 PR
 XX (UABR-) UAB RES FOUND.
 PA (UYFL) UNIV FLORIDA.
 XX
 XX Han N, Lantz M, Lepine G, Patti JM, Progulskie-Fox A;
 PI Tumwasorn S;
 XX WPI; 1996-287181/29.
 DR N-PSDB; AAT30645.
 DR
 XX Porphyromonas gingivalis genes and proteins - used in the detection
 PT and vaccination against periodontal disease
 PS Claim 4; Page 103-104; 153pp; English.
 XX
 XX Harep1 (AAR96021) is the product of the Harep1 repeat unit (AAT30645)
 CC of the hga gene (AAT30654) of P. gingivalis 318. It forms part
 CC of haemagglutinin hgaA (see also AAR96030). Harep1 and other
 CC hgaA repeat unit products (see also AAR96022-24) can be obtd. from
 CC transformed host cells and used as vaccines to protect humans or
 CC

CC incorporating AG-1.

XX Sequence 1704 AA;

Alignment Scores:

Pred. No.: 2,92e-77 Length: 1704
 Score: 716.00 Matches: 131
 Percent Similarity: 99.25% Conservatives: 2
 Best Local Similarity: 97.76% Mismatches: 1
 Query Match: 97.95% Indels: 0
 DB: 16 Gaps: 0

US-09-980-370-5 (1-402) x AAW34843 (1-1704)

QY 1 GCAGACTTCACGGAACGTTCCGAGTCTTCTACTCATGGAGAGCCAGCGGATGGACT 60
 DB 1137 AlaaspPheThrGluThrPheGluSerThrHisGlyGluAlaProAlaGluTrpThr 1156
 QY 61 ACTATCGATGCCGATGGCGATGGTGGAGGTTGGCTCTGTCTCTCCGACCAATTCGAC 120
 DB 1157 ThrileAspAlaaspGlyGlnGlyTrpLeuGlySerSerGlyGlnLeuAsp 1176
 QY 121 TGGCTCAGAGTCATGGCGGACACGAGTAGTAAGCTCTTCTCATGGAATGGCT 180
 DB 1177 TrpLeuThrAlaHisGlyGlyThrAsnValValAlaSerPheSerTrpAsnGlyMetAla 1196
 QY 181 TTGAATCCTGATACTATCTCATCTCAAGAGGTTTACAGCGCCAGCAAGGTAAGTAC 240
 DB 1197 LeuAsnProAspAsnTyrlleuSerLysAspValThrGlyAlaThrLysValLysTr 1216
 QY 241 TACTATCCAGTCAACGAGCTTCCGCGGATCACTATCGGTCGATGATCTCCAGACG 300
 DB 1217 TyrTyrlaValaAspGlyPheProGlyAspHisTyrlaValMetileSerLysThr 1236
 QY 301 GGCAGCAACGCGGAGACTTCACGGTTGTTTCGAGAAACGCTTACCGAATAAAG 360
 DB 1237 GlyThrAsnAlaGlyAspPheThrValValPheGluGluThrProAsnGlyIleAsnLys 1256
 QY 361 GCGGAGCAAGATTCGGCTTTCCACGGAAGCAATGGCGCC 402
 DB 1257 GlyGlyAlaArgPheGlyLeuSerThrGluAlaAsnGlyAla 1270

RESULT 5

AAW34843

ID AAW34843 standard; protein; 1704 AA.

XX AC AAW34843;

XX DT 03-JUN-1998 (first entry)

XX DE Arg-gingipain high molecular weight prepolyprotein sequence.

XX KW Arg-specific gingipain protease; gingivalis; periodontal disease;

XX KW vaccine; infection.

XX OS Porphyromonas gingivalis.

XX FH Key Location/Qualifiers

XX FT Protein 1..227

XX FT /note= "precursor protein"

XX PN W09734629-1.

XX PD 25-SEP-1997.

XX PF -21-MAR-1997; 97WO-US04635.

XX PR 22-MAR-1996; 96US-0013945.

XX PA (MORE-) MOREHOUSE SCHOOL MEDICINE.

XX PA (UYGE-) UNIV GEORGIA RES FOUND INC.

XX PI Genco CA, Potempa J, Travis J, Genco C;

XX

DR WPI; 1997-479993/44.

XX N-PSDB; AAT93872.

XX

PT Porphyromonas gingivalis Arg-specific gingipain protease peptide(s)
 PT - useful for protecting animals and humans from gingivalis and
 PT periodontal diseases

XX

PS Disclosure; Pages 68-73; 95pp; English.

XX

CC The present sequence represents an arginine-specific protease of
 CC Porphyromonas gingivalis. The following peptides, derived from Arg-
 CC and Lys-specific high molecular weight proteases, offer protection
 CC against infection: YTYTVYRDGK IKEGLTATTE DDGVTATGNHE YCCKEYTAGS VSPKVC
 CC (I); YTPVEKONG RMIVAKKY (II); QLPFIFDVAC VNGDFLSMP CFEALMRAO
 CC (III); GEPNPQVPS NLTTATGOK VTLKWDAPSTK (IV); GNHEYCVPEVK YTAGVSPKVC
 CC KDVTY (V); RMEFMYEPGR YTPVEKONG (VI); TFAGEDTYK RMEFMYEPGR (VII);
 CC DYTYTVYRDG TIKEGLTAT TFEEDGVATG NMEYCVCKVKY TAGVSPKVC (VIII); YTYTVYRDGT
 CC KIKEGLTATTE EEDG (IX); RDKIKIEGL TATTFEEDGV ATGN (X); KIKEGLTATT
 CC FEEDGVATGN HEY (XI); KWDAPNGTPN PNPEN PGTTLSE (XII); and
 CC YTPVEKONG RMIVAKKY (XIII). They are used in vaccines to protect
 CC animals, including humans, from gingivitis and/or periodontal
 CC diseases.

XX

SQ Sequence 1704 AA;

Alignment Scores:

Pred. No.: 2,92e-77 Length: 1704
 Score: 716.00 Matches: 131
 Percent Similarity: 99.25% Conservatives: 2
 Best Local Similarity: 97.76% Mismatches: 1
 Query Match: 97.95% Indels: 0
 DB: 18 Gaps: 0

US-09-980-370-5 (1-402) x AAW34843 (1-1704)

QY 1 GCAGACTTCACGGAACGTTCCGAGTCTTCTACTCATGGAGAGCCAGCGGATGGACT 60
 DB 1137 AlaaspPheThrGluThrPheGluSerThrHisGlyGluAlaProAlaGluTrpThr 1156
 QY 61 ACTATCGATGCCGATGGCGATGGTGGAGGTTGGCTCTGTCTCTCCGACCAATTCGAC 120
 DB 1157 ThrileAspAlaaspGlyGlnGlyTrpLeuGlySerSerGlyGlnLeuAsp 1176
 QY 121 TGGCTCAGAGTCATGGCGGACCAACGTAAGCTCTTCTCATGGAATGGCT 180
 DB 1177 TrpLeuThrAlaHisGlyGlyThrAsnValValAlaSerPheSerTrpAsnGlyMetAla 1196
 QY 181 TTGAATCCTGATACTATCTCATCTCAAGAGGTTTACAGCGCCAGCAAGGTAAGTAC 240
 DB 1197 LeuAsnProAspAsnTyrlleuSerLysAspValThrGlyAlaThrLysValLysTr 1216
 QY 241 TACTATCCAGTCAACGAGCTTCCGCGGATCACTATCGGTCGATGATCTCCAGACG 300
 DB 1217 TyrTyrlaValaAspGlyPheProGlyAspHisTyrlaValMetileSerLysThr 1236
 QY 301 GGCAGCAACGCGGAGACTTCCACGGTTGTTTCGAGAAACGCTTACCGAATAAAG 360
 DB 1237 GlyThrAsnAlaGlyAspPheThrValValPheGluGluThrProAsnGlyIleAsnLys 1256
 QY 361 GCGGAGCAAGATTCGGCTTTCCACGGAAGCAATGGCGCC 402
 DB 1257 GlyGlyAlaArgPheGlyLeuSerThrGluAlaAsnGlyAla 1270

RESULT 6

AAW67396

ID AAW67396 standard; Protein; 1704 AA.

XX AC AAW67396;

XX DT 25-APR-2000 (first entry)

XX DE Arg-gingipain-2 amino acid sequence.

